

School of Electrical Engineering, Computing and Mathematical
Sciences

**Simulating Radio Variability Using Tracers of the
Interstellar Medium**

Elliott George Charlton

0000-0002-8000-1171

This thesis is presented for the Degree of
Master of Research (Physics)
of
Curtin University

November 2021

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Signature:

Date: 03/11/2021

Statement of Contribution by Others

To Whom It May Concern,

I, Elliott George Charlton, wrote this thesis. My lead supervisor Dr Paul Hancock created the idea and base program for the project, the SM2017 model. The RISS19 model described in this work was derived from the SM2017 model, with contributions from my self and Dr Hancock. The simulation code that relies on the RISS19 model is entirely of my own creation.

Dr Paul Hancock and A/Prof Jean-Pierre Macquart had input in the technical process I created for this work. They, along with Dr John Morgan, had input in the construction and editing of the thesis.

Candidate Signature:

Supervisor Signature:

Date: 03/11/2021

Abstract

Variability, a change in brightness over time, is observed in the majority of radio sources. The origin of this variability is dependent on several factors such as the observing frequency and size of the source. For the most common compact radio sources such as active galactic nuclei (AGN), the dominant cause of variability is interstellar scintillation (ISS), specifically Refractive ISS (RISS) at sub gigahertz frequencies, due to the interstellar medium (ISM). RISS induced variability has a significant effect on the interpretation of gigahertz radio observations of compact sources on timescales of hours to years.

To date several models exist that map the ISM in order to predict Diffractive ISS (DISS) for pulsar and fast radio burst studies. However, no such model exists for studies of extra-galactic radio sources for which RISS is the dominant cause of variability. To fill this gap, the software SM2017 ([Hancock et al., 2019](#)) was created to predict RISS using Hydrogen alpha ($H\alpha$) measurements as a proxy for electron density. We build upon the SM2017 model to create an updated model: RISS19 which is able to predict the magnitude and incidence of scintillation induced radio variability for a given radio survey. In this work we compare predictions of the RISS19 model to observations from the High Galactic Latitude (HGL, [Hancock et al. \(2019\)](#)) survey carried out with the Murchison Widefield Array (MWA) at 185MHz. Using RISS19 in combination with scattering theory, source size, and spatial distributions, we can simulate regions of sky and produce population statistics for the radio variability an observer would be expected to see from any given area of the sky. The model can then be used to determine the scintillation contribution of individual sources or entire survey regions.

In our results we were able to use the RISS19 model to predict variability for line of sights, as well as limiting the variability found in an SED. A simulation is created to test how well the RISS19 model predicts population statistics, as well as our ability to generate a realistic source population as seen by a telescope. In

order to verify the predictions of the RISS19 model, eight different surveys were selected of varying frequency (185, 1400, and 5500 MHz), location, area, and sensitivity. It is found that the simulations prediction of variable and non-variable sources is bound by the input fraction of point sources we use this to determine the number of point sources required to match observed survey results. We group the surveys into their frequency bands and find that we require 0.09, 3.51, and 19.61% of the population to be point sources at 185, 1400, and 5500 MHz respectively. Further, we constrain the source size of these populations based on the critical angular scale of each survey area: 16.3, 0.053, and 0.0037 milli-arcseconds at 185, 1400, and 5500 MHz respectively.

The code used within this thesis is can be found at github.com/PaulHancock/RISS19.

Acknowledgements

I would like to thank my supervisor, Dr. Paul Hancock, and express my sincere appreciation for his support, mentoring, and expertise, not just throughout this project but throughout the three years I have been his student. I would not be the person I have become without Dr Hancock's encouragement and guidance throughout my learning and research career. Academically I could not have asked for a better mentor academically and for me, there could not have been a better role model.

I would also like to thank my co-supervisors. Firstly, I would like to thank the late Associate Professor Jean-Pierre Macquart for his expertise, observations and critique throughout the project, he was a brilliant man who will be missed dearly by the community. Secondly, I would like to thank Dr. John Morgan for his fantastic feedback and assistance on the project.

My thanks too to Curtin University, CIRA and the entire staff, for all of their resources, guidance and support that they have provided to myself and fellow students over the years.

Thanks also to my family and friends; thank you for supporting me throughout this journey. Thank you for always being there for me and for being the very best of companions over the years. Furthermore I would like to especially like to thank my love, Jess. You have always been there for me and I could not have completed this journey without your love, happiness, encouragement and your smile. Thank you for being you.

Finally a big thank you to my friends, colleagues and fellow students at CIRA with whom I have spent countless hours studying and working with. I particularly wish to thank Ben, Cath, Jaiden, Kat, Pikky, and Ronniy; thank you for keeping me sane over throughout this project, thank you for your guidance, help, laughs and good company.

Contents

1	Background	1
1.1	Introduction	1
1.2	Observational Radio Astronomy	4
1.2.1	Variability	7
1.2.2	Variable Radio Source Population	9
1.2.2.1	Stellar Remnants	9
1.2.2.2	Pulsars and Fast Radio Bursts	10
1.2.2.3	Galaxies	12
1.2.3	Active Galactic Nuclei	13
1.2.4	Doppler Boosting	16
1.2.5	Intra-day Variability	18
1.3	The Intervening Medium	20
1.3.1	The Interstellar Medium	21
1.3.2	Mapping Turbulence and Inhomogeneities	23
1.3.2.1	Density Fluctuation Relations	23
1.3.2.2	Dispersion Measure	24
1.3.2.3	Hydrogen Alpha	26
1.4	Scintillation and Scattering	27
1.4.1	Weak Scattering	31
1.4.2	Strong Scattering	33
1.4.2.1	Diffraction Scintillation	33
1.4.2.2	Refractive Scintillation	34
1.4.3	Application	35
1.5	Summary	37

2	Scintillation and Scattering Models	39
2.1	Introduction	39
2.2	SM2017	40
2.3	RISS19	41
2.3.1	Results from RISS19	43
2.4	Improvements to RISS19	48
2.4.1	Hydrogen Alpha Maps	48
2.4.2	Distance Models	50
2.5	Results	56
2.5.1	Comparison of Distance Models	57
2.5.2	Comparison to HGL	58
2.5.3	Multi-Frequency Data	59
2.6	Summary	61
3	Simulation and Surveys	63
3.1	Introduction	63
3.2	Position Generation	64
3.3	Source Counts	65
3.4	Source Structure	68
3.4.1	Source Structure and Size Distribution	70
3.5	Variability Calculation	76
3.6	The Simulation	78
3.7	Results and Discussion	83
3.7.1	Comparison to HGL	84
3.7.2	Comparison to Surveys	88
3.7.3	Current State	92
4	Conclusion	95

List of Figures

1.1	The galaxy Centaurus A. Wavelengths and telescopes displayed in each panel from top left to bottom right: visible light (La Silla); infrared (Spitzer); millimetre/submillimetre and visible light (ALMA); radio (VLA); x-ray, radio and visible light; x-rays (Chandra). Source: ESO ¹	3
1.2	Different wavelengths are absorbed or transmitted by the atmosphere. Two main wavelength ranges can be seen from ground based telescopes: optical and radio. Source: NASA ²	5
1.3	The radio source J0034.0–6639, classified first by Saripalli et al. (2012) as a low surface-brightness, restarted radio galaxy. The figure demonstrates how observing parameters such as frequency, sensitivity and angular resolution impact the science which can be conducted. From left to right the top row panels show: GLEAM 200 MHz, ATLAS (low-res) 1.4 GHz, with ‘North’ (N) and ‘South’ (S) lobes of the galaxy labelled, WISE infrared. From left to right the bottom row of panels show: SUMSS 843 MHz, ATLAS (high-res) 1.4 GHz, and the spectral energy distribution of the source with points from GLEAM (green) and ATLAS. Image compiled using code from Benjamin Quici.	8

- 1.4 Transient parameter space of specific luminosity versus the product of observing frequency and transient duration. This figure shows the peak luminosity for several different groups of transient events. It also shows the divide between coherent and incoherent emission, lines of equal brightness temperature is also shown (dashed lines). The most relevant to this work are the AGN/blazar/QSO and the pulsars/fast radio bursts groupings. Source: Pietka et al. (2014). 10
- 1.5 **(a)** the euclidean normalised 1.4 GHz source counts from Rawlings et al. (2015), where the source counts are the number of sources observed at a given flux. Relevant colours and symbols as follows: total, blue squares; SF-powered galaxies, black stars; AGN-powered galaxies, red circles; AGN/SF hybrid, orange diamonds. **(b)** the 150 MHz luminosity function from Sabater et al. (2019). Relevant colours and symbols as follows: SF galaxies (blue diamonds); AGN (red circles); and the 1.4 GHz converted luminosity function (dashed lines). 13
- 1.6 The *unification* model for radio galaxies, which divides the types of radio galaxies into radio- loud and quiet depending on the magnitude of their luminosities. Radio galaxies are then further broken down by the viewing angle of the observer. Also shown is the structure and emission sites of these galaxies. Source: FERMI³ 14
- 1.7 Radio galaxy PKS 2250-351 at a redshift of $z=0.2115$ observed using the Australian Square Kilometre Array Pathfinder (ASKAP) at a radio frequency of 887 MHz, the angular size of the source is 5.92 arc minutes. Labelled in this figure are the typical features of radio galaxies such as the core, jet, lobe and hotspot. Note that the counter-jet is not visible due to the orientation of the galaxy. More details on this source can be found at Seymour et al. (2020). Image compiled using code from Benjamin Quici. 15

1.8 Comparison between (a) a non-Doppler boosted and (b) Doppler boosted source observed using 15 GHz images with VLBA from Kellermann et al. (2007). (a) The radio galaxy NGC1052 is almost perpendicular to our line of sight and so the jets have little to no Doppler boosting and look symmetric. The tick marks are two milli-arcseconds. (b) The Quasar PKS 1148-001 jets are Doppler boosted and the intensity of the observed radiation will be brighter and dimmer for the jet and counter jet respectively. The counter jet is not visible in this image. The tick marks are one milli-arcsecond. 17

1.9 Some of the different media intersecting our line of sight, as a function of increasing distance. From closest (bottom) to furthest (top) the mediums are: ionospheric, interplanetary, interstellar, and intergalactic. 22

1.10 Pulse broadening over frequency. The diagram shows pulsar B1831-03's pulse profile at five different frequencies. The diagram emphasises the effect of frequency on the scattering and shows that at lower frequencies it has a significant effect on the pulse profile (also shown in Equation 6). The solid line is the exponential model fit for the data. Lorimer & Kramer (2005) 26

1.11 Hydrogen alpha map created by Finkbeiner (2003), measured in Rayleighs, displayed in galactic coordinates. 27

1.12 This figure has been recreated from Narayan et al. (1992), displays the regimes of weak and strong (refractive and diffractive) scattering. It shows the change in modulation as a function of the lengthscales r_F and r_{diff} as well as the frequency. The green band shows the frequency range for the MWA (70-300 MHz) and the red line shows the transition frequency of approximately 5 GHz for this line of sight. 28

1.13 Thin screen model. Figure shows the effect of an inhomogeneous medium collapsed into a thin screen has on a wave when it passes through it. 29

1.14 Representation of the first Fresnel zone on the scattering screen. The Fresnel scale describes effects seen from purely geometric time delay. A source’s signal can be received from any position on the plane. Within the first Fresnel zone of size r_F these signals add coherently. Diagram created by Jessica Wu (2019). 31

1.15 Representation of diffractive lengthscale, on the scattering screen. r_{diff} represents zones where the phase change due to inhomogeneities is less than one radian and is therefore coherent. These patches have approximate size r_{diff} as well as being spaced on average by r_{diff} . Diagram created by Jessica Wu (2019). 32

1.16 Coherent patches of r_{diff} act like slits and re-emit cones of light into angle θ_{scatt} . An observer sees patches within r_{ref} . Diagram created by Jessica Wu (2019). 34

1.17 Cones of radiation emitted by the diffractive patches can be focused/defocused by a large scale phase front, occurring on scale r_{ref} . Where an observer would view more/less patches depending on whether the phase front is being focused towards or defocused away from an observer. The above diagram shows an example of diffractive patches being focused by a lens of size r_{ref} . Diagram created by Jessica Wu (2019). 36

2.1 (a) SM2017 and (b) RISS19 program flowcharts. Green boxes input data and the dotted connectors represent parameters which have a default value but can be changed manually if desired. Yellow, blue and red boxes show calculated outputs where blue and red can be outputted into an output table. red boxes indicate new parameter calculations included within RISS19. The scattering disk parameter is only considered in other calculations when a source size is given to the program initially. 44

2.2	Figure showing the average total modulation for a point source (a) and the modulation expected on a one year period for a point source (b). The three different frequencies displayed are: 100 MHz (blue), 185 MHz (orange) and 1 GHz (green). The solid line represents the average value and the shaded cells are the standard deviation. Figures from Hancock et al. (2019).	45
2.3	Figure showing the timescale of modulation of refractive scintillation for each latitude at each frequency. The three different frequencies displayed are: 100 MHz (blue), 185 MHz (orange) and 1 GHz (green). The solid line represents the average value and the shaded cells are the standard deviation. Figure from Hancock et al. (2019).	46
2.4	Variability versus $H\alpha$. This plot shows the HGL measured data (blue triangles) and the trend of reduction in variability at higher $H\alpha$ as predicted by RISS19 (red circles). Figure from Hancock et al. (2019).	47
2.5	Left: comparison of the averaged (over the longitude) $H\alpha$ distributions with the variance shown as the shaded parts of each line. Right: The distribution of average (over the longitude) fractional errors in the $H\alpha$ measurements. In both plots the orange is the All-sky WHAM map. and blue Finkbeiner map. Adapted from Hancock et al. (2019).	50
2.6	(a) Shows the results from Hancock et al. (2019). (b) Plot showing the results from Hancock et al. (2019) recalculated using the new $H\alpha$ map.	51
2.7	The percentage error of the Finkbeiner (top) and AS-WHAM (bottom) Map. The flux scale ranges from 0.1% (black) to 100% error (white).	52
2.8	The $H\alpha$ maps from the Finkbeiner (top) and AS-WHAM (bottom) data. The flux scale ranges from 1 Rayleigh (black) to 316 Rayleigh (white). Due to the authors selection of velocity ranges, the Magellanic Cloud is lost in the AS-WHAM data.	53

2.9 Figures showing the distribution of n_e (black) down a line of sight, with three different models for weighting the screen distance. The different distance models are: 1 kpc (blue), ‘pancake’ (green) and the weighted fit (red). The 3 different models are shown for 4 lines of sight. 54

2.10 The Bhat relation (Bhat et al., 2004). Here we can see the pulse delay in milliseconds versus the dispersion measure. Filled circles are measurements made by Bhat et al. (2004), crossed circles are derived from decorrelation bandwidths and open circles are other published pulsar measurements. The solid line represents the best-fit model for this relation. 55

2.11 Plot of the pulsar (blue), FRB (red) and other galactic components (green) used to generate the YMW16 model, plotted in Galactic coordinates as an aitoff projection. Data obtained from Yao et al. (2017). 56

2.12 The map of the pulse delay in seconds, plotted as a function of Galactic latitude (l) and longitude (b). 56

2.13 Figures comparing each distance model: **(a)** point source modulation **(b)** scaled modulation on 1 year timescales. The three different models displayed are: RISS19 (blue), WD-RISS19 (orange) and TAU19 (green). The solid line represents the average value and the shaded cells are the standard deviation. Figures adapted from Hancock et al. (2019). 58

2.14 Comparison of the expected modulation of the different scintillation distance models against observational variability data from the HGL Survey (blue circles). RISS19 model is described by orange triangles, and the predictions of TAU19 model by green triangles.. Figure adapted from Hancock et al. (2019). 59

- 2.15 **(a)** Comparison of the expected modulation of the different scintillation distance models against observational variability data from the HGL Survey (blue circles). The predictions of the RISS19 model at its full modulation are shown by orange triangles, and at 448 days by magenta squares. The predictions of TAU19 model are shown by green triangles, and at 448 days by cyan squares. **(b)** The change in variability between the 448 day predicted modulation (square markers) and the maximum modulation (triangle markers), RISS19 is shown in orange and TAU19 is shown in green. Figure adapted from Hancock et al. (2019). 60
- 2.16 Top: Spectral energy distribution for one of Ms. Ross' sources. This diagram shows the log flux density (mJy) versus the frequency (MHz). In blue is the first year data, in green is the second year data. Further in blue shaded areas we have the predicted modulation of flux from the RISS19. Bottom: Residuals for the SED, the data shown is: zero line (dot-dashed black line), year 2 - year 1 (orange circles) and the modulation predicted by RISS19 (blue shaded area). 61
- 3.1 Top: The Euclidean normalised differential source counts for 150-154 MHz, where the black circles are Franzen et al. (2019), cyan circles Intema et al. (2017), pink squares Hales et al. (2007), and blue squares McGilchrist et al. (1990). Bottom shows the residuals between GLEAM (Hurley-Walker et al., 2016) and TGSS (Intema et al., 2017). Figure taken from Franzen et al. (2019). 66
- 3.2 The differential source counts for each survey at their respective frequencies. In blue is Franzen et al. (2019) at 154 MHz and in red is Hopkins et al. (2003) at 1400 MHz. 67
- 3.3 Figure showing the size scattering disk size (in milli-arcseconds) at different frequencies and latitudes. Data shows, 100 MHz (blue), 185 MHz (orange) and 1 GHz (green). The solid line represents the average value and the shaded cells are the standard deviation. Figure adapted from Hancock et al. (2019). 69

3.4	The simulation program flowchart, where green boxes are the inputs from the user, blue boxes are the simulation outputs, and the purple box is the RISS19 output.	80
3.5	Figure showing the number of sources versus the flux found the HGL survey (blue), the simulation (orange) and the simulation with an estimated completeness limit implemented (green). The 5 sigma cutoff is also shown as a red line.	85
3.6	The flux in Jansky of the local rms found in the HGL survey. . . .	86
3.7	The variability versus $H\alpha$ intensity for the HGL survey. The data shown: HGL Survey (blue circles), RISS19 model at its max modulation (orange triangles), and the green region simulated for point sources (upper limit) and 37% Point sources (lower limit). Figure adapted from Hancock et al. (2019).	88
3.8	Surface density for each survey (black crosses), the simulations for 100% point source population (blue squares), 37% population (orange squares) as seen in Table 3.7. Bars with triangles indicate upper limits where less than one variable source was detected within the survey. The plot has been separated into three frequencies, 185 MHz (green), 1400 MHz (red), and 5.5 GHz (purple). For each frequency the factor of point sources required to match the target survey has been plotted as: 19.61% for 5.5 GHz (lime plus), for 1400 MHz 3.51% (teal plus), and for 185 MHz 0.09% (red plus).	91

Chapter 1

Background

1.1 Introduction

From the first people gazing up at the stars in the night sky to detecting gravitational waves, astronomy has been a tool to wonder and learn about the Universe. The human eye is limited in both resolution (1 arcminute) and wavelength (380-750 nm), but with the development of new technology and better telescopes, astronomers have been able to open new windows into the Universe. Now, astronomers use the full electromagnetic spectrum: from short-wavelength gamma rays to long-wavelength radio waves. Observing at different wavelengths allows astronomers to look at the Universe from many perspectives, with each view providing a unique piece of the puzzle. For example in [Figure 1.1](#) we can see 6 different images of a nearby galaxy - Centaurus A. In the visible spectrum we can see the light produced by all of the stars, the infra-red radiation allows us to see the dusty stellar disk, radiation at sub-millimetre/millimetre wavelengths allow us to view the distribution of molecules, investigate dense dark clouds of gas and explore cold matter, radiation at \sim meter wavelengths allows us to see the plasma launched away from a central super-massive black hole, and finally x-ray radiation allow us to observe the highest-energy regions within the galaxy such as the core and inner jets. Multi-wavelength studies of objects are clearly advantageous as they provide access to different types of information from each wavelength, and it is clear to see from [Figure 1.1](#) that we have a more complete picture when we combine wavelengths together.

One of the many methods by which an astronomer studies space and the sources within it, is to conduct astronomical surveys. Astronomers have employed ground- and space-based telescopes to develop all-sky surveys which observe the Universe at different wavelength regimes. Surveys of the sky allow astronomers to collect information about the population of sources in the observed region, which are then used to create catalogues. A catalogue is a collection of information on the observed sources, which can include but is not limited to the: position, spectra, polarisation, and brightness. Catalogues can be used to infer trends and information about sources.

To understand more about the evolution of the Universe and the sources within it we must look outside of our Galaxy to the multitude of extra-galactic sources – e.g. galaxies. These sources can show astronomers what young and old galaxies look like and how they evolve. Surveying extra-galactic sources provides new challenges for astronomers, and while surveys in different wavelength regimes contribute unique information, they also have unique limitations. Unlike optical wavelengths, radio waves do not suffer the same obscuration by dust, and therefore allow astronomers to peer deeper into the Universe.

The radio emission observed from galaxies can be broadly grouped into two categories describing their origin; the radio emission produced via stellar processes within the galaxy disk, and, the radio emission produced via the accretion of matter onto the super-massive black hole at the galaxy centre. While these categories display different properties (see [Section 1.2](#)), for now it is important to recognise that typically this emission evolves over very long timescales (Myr–Gyr), meaning these radio sources can effectively be treated as ‘static’ with respect to a characteristic human observing timescale.

Although most objects would be considered static, there is a population of

¹[www.almaobservatory.org/en/about-alma-at-first-glance/how-alma-works/
how-does-alma-see/](http://www.almaobservatory.org/en/about-alma-at-first-glance/how-alma-works/how-does-alma-see/)

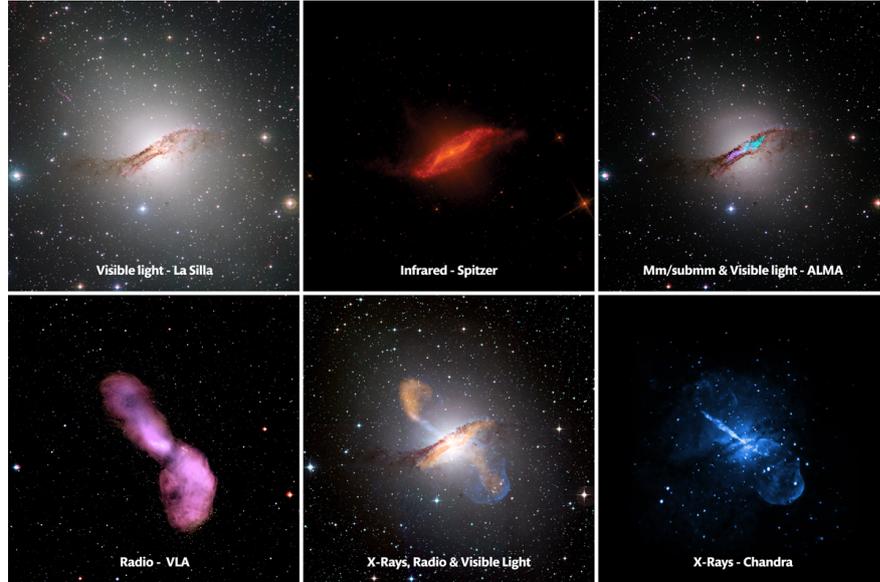


Figure 1.1: The galaxy Centaurus A. Wavelengths and telescopes displayed in each panel from top left to bottom right: visible light (La Silla); infrared (Spitzer); millimetre/submillimetre and visible light (ALMA); radio (VLA); x-ray, radio and visible light; x-rays (Chandra). Source: ESO¹.

sources that vary over time. Studying objects whose brightness varies with time are of particular interest as this variability allows us to analyse the astrophysics of their emission. Variability is often categorised into either: intrinsic or extrinsic variability. These categories are designed to reflect differing origins for the observed variability. For example, variations occurring interior to the source will give rise to intrinsic variability, otherwise if the variation occurs exterior to the source it is extrinsically variable. Many astronomical objects show one of the two previously-mentioned forms of variability; from objects in our own Galaxy such as flaring stars or rotating neutron stars (pulsars), all the way to the scintillation of distant galaxies. As will become evident, variability forms a key component of astrophysical processes within our Universe and therefore understanding the how, what and why of variability is important to astronomers.

Over time astronomers have gained a greater understanding of the mechanisms and causes of variability; this is particularly relevant to our understanding of variability at radio wavelengths. Historically, it was assumed that variability at radio wavelengths arose intrinsically, however this perception was changed when

[Hunstead \(1972\)](#) closely examined four radio sources to show alternative evidence in favour of extrinsic variability. Today, it is widely accepted that much radio variability can be largely attributed to the propagation of radio waves through an intervening medium. However, as variability is an observed effect it is still quite difficult to disentangle and correctly attribute what is intrinsically or extrinsically varying without a closer examination of the source. Variability of the radio sky is simultaneously a tool for astronomers to learn about sources, and a hindrance for calibrations. As such, predicting the extrinsic radio variability expected within a region of sky becomes an important exercise.

It is the goal of this thesis to build a tool that will allow astronomers to predict extrinsic radio variability due to the intervening media. Some efforts have already been made to predict extrinsic radio variability by Dr. Paul Hancock using his SM2017 model which uses measurements of hydrogen alpha ($H\alpha$) in the galactic plane to predict the magnitude of variability. Our goal will be to further build upon this work by updating and improving the SM2017 model. Further, we will develop a simulation to generate a population of sources to predict the variability in any part of the radio sky. This work will be compared to the *High Galactic Latitude* (HGL) survey conducted by [Hancock et al. \(2019\)](#) using the Murchison Widefield Array (MWA) at 185 MHz.

1.2 Observational Radio Astronomy

Since the detection of the first radio signal originating from outside Earth in the early 1930s by Karl Jansky ([Jansky, 1933](#)), radio astronomy has played a key role in the advancement of our knowledge of the Universe. Before the invention of satellites, optical and radio astronomy were the prevalent branches of astronomy as these wavelengths were not significantly blocked by the Earth's atmosphere ([Figure 1.2](#)). Radio wavelengths range from tens of kilometres to approximately one millimeter (~ 300 GHz), with the longest wavelength end of the spectrum being blocked by the Earth's ionosphere.

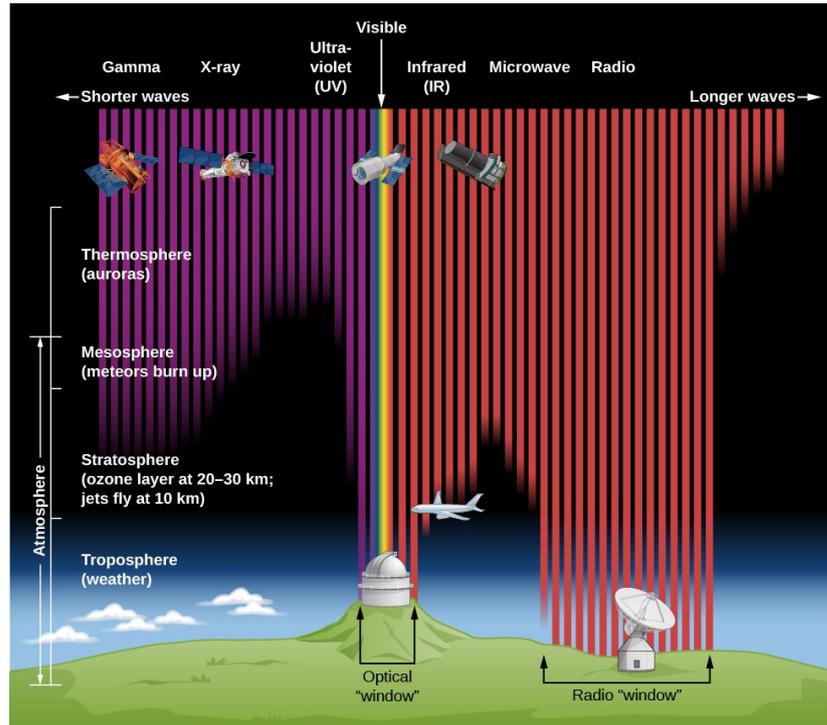


Figure 1.2: Different wavelengths are absorbed or transmitted by the atmosphere. Two main wavelength ranges can be seen from ground based telescopes: optical and radio. Source: NASA²

The radio regime is often observed with a radio interferometer - an array of telescopes separated by some distance (baselines) that work together to collect signals. The field of view and angular resolution of a radio interferometer are dependant on the observing frequency (λ). The field of view (Ω_{FOV}) is also determined by the size of a single element (D):

$$\Omega_{FOV} \approx \frac{\lambda}{D}, \quad (1.1)$$

whereas the angular resolution (θ) for interferometers is determined by the largest baseline separating observing elements (B_{max}):

$$\theta \approx \frac{\lambda}{B_{max}}. \quad (1.2)$$

Astronomers use telescopes by undertaking pre-planned and coordinated observations of an object or part of the sky (surveys). The design of a radio survey

²imagine.gsfc.nasa.gov/science/toolbox/emspectrum1.html

is often motivated by a set of key science goals that the survey will explore. In practice, these research goals are limited by the constraints imposed by the telescope itself; observing frequency, instrumental sensitivity, angular and spectral resolution, and field of view. Physical limitations, as well as practical limitations such as time and funding, often mean that no telescope will be able to satisfy each of these requirements, and by extension implies that all radio surveys will have some technical limitation. For example, the large field of view and excellent low-frequency coverage of the MWA makes this an excellent instrument to survey the global spectral properties of radio sources. However, studying the intricate, small-scale structure of radio sources is not feasible with the MWA due to its low angular resolution. By combining surveys that excel in different parameter spaces we can gain more information about sources found within each survey.

To date there are a large number of radio surveys, we have compiled a few notable surveys to compare and demonstrate how each survey can be combined to build a complete picture of a source. Among these surveys are: GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) - [Hurley-Walker et al. \(2016\)](#), Sydney University Molonglo Sky Survey (SUMSS) [Mauch et al. \(2003\)](#), and the Australia Telescope Low-Brightness Survey (ATLBS) [Subrahmanyan et al. \(2010\)](#). Where the parameters of each survey are shown in [Table 1.1](#).

Survey Catalogue	Frequency (MHz)	Sensitivity (mJy/beam)	Angular Res (arcsec)	Sky Area (sq deg)
GLEAM	151	6-10	100	30,000
SUMSS	843	1	45	8,100
ATLBS (low-res)	1400	0.08	50	8.42
ATLBS (high-res)	1400	0.08	6	8.42

Table 1.1: General parameters of example surveys : GLEAM [Hurley-Walker et al. \(2016\)](#), SUMSS [Mauch et al. \(2003\)](#), and ATLBS (low/high-res) [Subrahmanyan et al. \(2010\)](#).

In [Figure 1.3](#) we see the compiled observations of radio galaxy J0034.0–6639 at several different frequencies, resolutions and sensitivities. The radio spectrum arising from the lobes encodes the energetics and evolutionary history of the radio source; adequately sampling the radio spectrum requires access to a range of observing frequencies. The GLEAM survey provides excellent frequency coverage over 72–231 MHz, however, its low angular-resolution prohibits the study of any fine radio structure. The increased resolution of the ATLAS (low-resolution) survey reveals structure within the lobes that is otherwise missed by GLEAM. At a fixed observing frequency, incorporating longer baselines is needed in order to improve the angular resolution ([Equation 1.2](#)). This often comes at a cost to the brightness sensitivity, meaning that diffuse emission is more easily missed. This can be seen in the comparison between the ATLAS low- and high- resolution surveys; while the lobes are easily detected at lower resolution, they become much more difficult to detect at high resolution.

By combining multiple survey catalogues we are able to learn about the global properties of the radio source population. We do this by examining sources in different parameter spaces, such as a sources brightness, as a function of frequency (spectral energy distribution) or as a function of time (light curve). In this thesis we focus on sources that change their brightness over time - variable sources.

1.2.1 Variability

Often the study of radio sources begins by measuring their brightness as a function of a specific domain, such as time, frequency or polarisation. Temporal changes in the brightness of radio sources gives rise to the study of variability. Further, radio variability can be used to infer information about a source or the environment that surrounds it. There are two classes of variability: intrinsic, and extrinsic. Intrinsic variability occurs when a source’s brightness is affected by an underlying mechanism within the source itself. Examples of this are: flaring stars or a change in the accretion state of black holes. Extrinsic variability describes a process where the intensity of an incoming signal is altered between

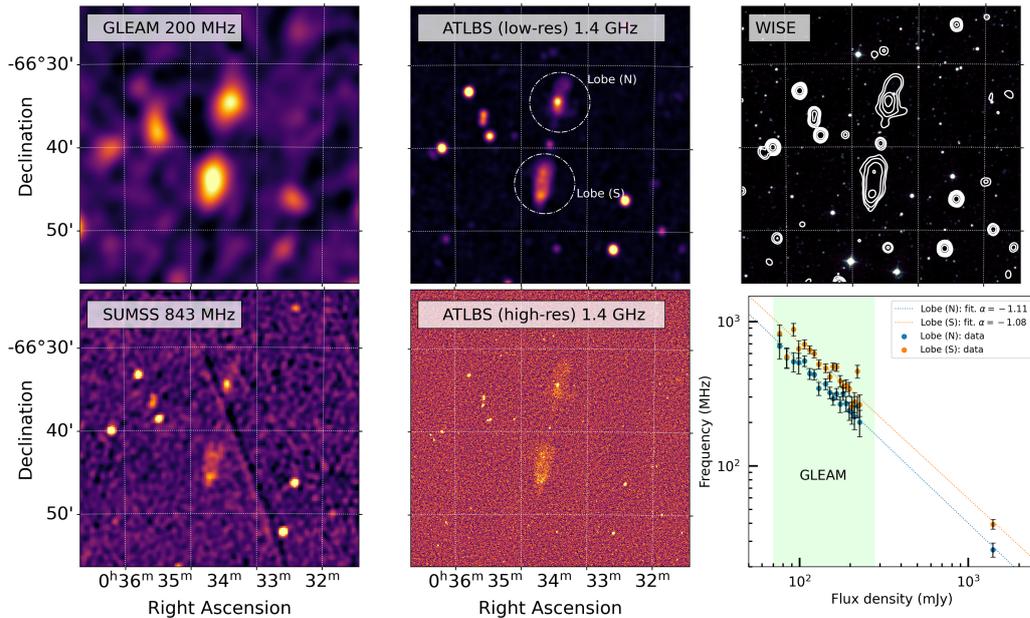


Figure 1.3: The radio source J0034.0–6639, classified first by Saripalli et al. (2012) as a low surface-brightness, restarted radio galaxy. The figure demonstrates how observing parameters such as frequency, sensitivity and angular resolution impact the science which can be conducted. From left to right the top row panels show: GLEAM 200 MHz, ATLBS (low-res) 1.4 GHz, with ‘North’ (N) and ‘South’ (S) lobes of the galaxy labelled, WISE infrared. From left to right the bottom row of panels show: SUMSS 843 MHz, ATLBS (high-res) 1.4 GHz, and the spectral energy distribution of the source with points from GLEAM (green) and ATLBS. Image compiled using code from Benjamin Quici.

the source and the observer. Extrinsic variability can be driven by scattering and scintillation, gravitational lensing or even a change in the absorption of an intervening gas. A major source of extrinsic variability at radio frequencies is Interstellar Scintillation (ISS) which is driven by turbulence in the Interstellar Medium (ISM).

The timescale on which an object’s observed intensity is changing, is fundamental in variability studies. Intrinsic variability encodes the timescale across which physical changes occur within a source, e.g. allowing one to probe its physical size and mechanisms related to their intrinsic variation. On the other hand, the timescales of extrinsic variability are driven by the intervening medium, and

can be used to measure properties such as the electron density along the line of sight. With a fundamental understanding of the causes of intrinsic and extrinsic variability, it is possible to disentangle these two effects, although this division is not always clear. Historically variability has been most notable in varying extra-galactic sources, such as compact active galaxies.

1.2.2 Variable Radio Source Population

When astronomers look out into the radio sky they are greeted by numerous different variable objects, both galactic and extra-galactic. These sources are categorised in [Figure 1.4](#) as a function of their peak luminosity during transient event and the product of observing frequency and transient duration. Some of the notable groupings are: x-ray binaries and supernovae (stellar remnants), AGN/blazar/QSO (galaxies), pulsars, and fast radio bursts. We discuss each of these in detail below.

1.2.2.1 Stellar Remnants

Stellar evolution ends with the formation of a stellar remnant through a supernova event. A supernova is a bright event, whereby the mass stored in the outer shells of a star are blown off via a highly-energetic explosion. There are several type and sub-types of supernovae, although they can be broadly grouped by cause of the supernovae: 'thermal runaway' (Type Ia) and 'core collapse' (Type Ib/c, Type IIa and IIb), all except Type Ia have been observed in the radio regime ([Lien et al., 2011](#)). Of the 50 radio counterparts of supernovae observed, these events have been found to have a typical duration of 10^2 days.

The stellar remnants formed by supernovae are also called compact stars, the most common being: white dwarfs, neutron stars and black holes. A compact star in a binary system can form an x-ray binary when the main-sequence star (donor) begins to accrete matter onto a black-hole or neutron star. X-ray binaries

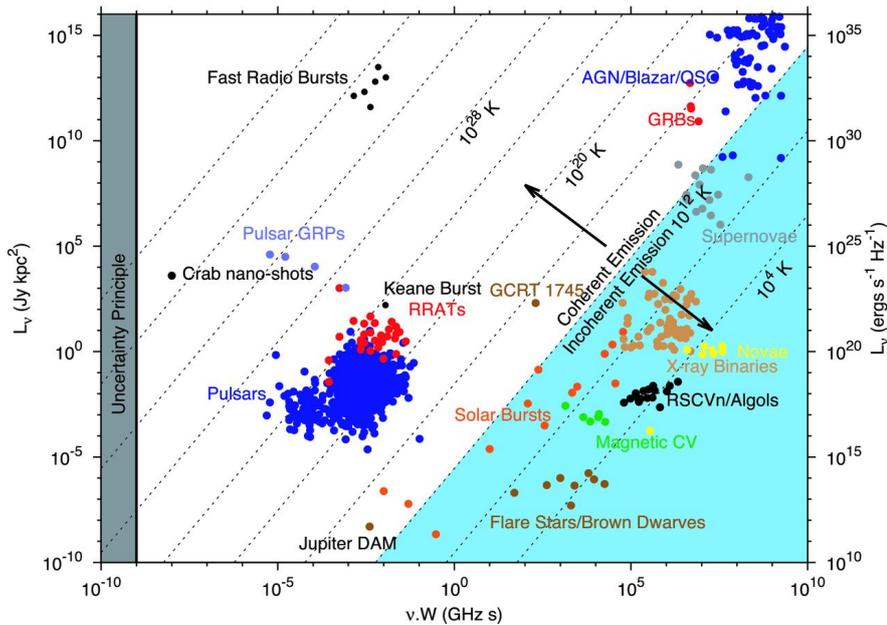


Figure 1.4: Transient parameter space of specific luminosity versus the product of observing frequency and transient duration. This figure shows the peak luminosity for several different groups of transient events. It also shows the divide between coherent and incoherent emission, lines of equal brightness temperature is also shown (dashed lines). The most relevant to this work are the AGN/blazar/QSO and the pulsars/fast radio bursts groupings. Source: [Pietka et al. \(2014\)](#).

are luminous in the X-ray regime but can also have a strong component at radio wavelengths due to the emission of synchrotron radiation from jets. Black-hole binaries have defined X-ray states (soft/hard) which are correlated with radio emission. In the hard state there is a radio emission from a collimated, flat (spectrum), and relativistic jet which is quenched in the soft state ([Fender et al., 2004](#); [Corbel et al., 2004](#)). Variability in these black-hole binaries arise during the transition between these states as the jet is quenched and then re-established, this can occur on short timescales of minutes to days ([Dhawan et al., 2000](#); [Bright et al., 2020](#)).

1.2.2.2 Pulsars and Fast Radio Bursts

A pulsar is a rapidly rotating neutron star that produces beams of radiation along its magnetic poles as it rotates. The neutron star is formed when a progenitor of $\sim 10 - 25$ solar masses collapses to form a remnant core of approximately

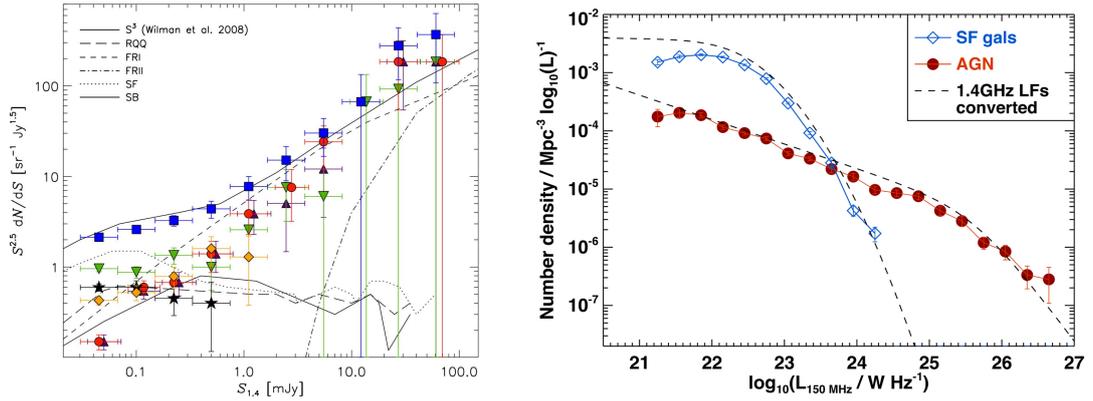
$\sim 1.1 - 2.1$ solar masses during a supernova explosion. The mass limit of neutron stars is known as the Tolman-Oppenheimer-Volkoff Limit, which has been further constrained over time using gravitational wave events (Kalogera & Baym, 1996; Margalit & Metzger, 2017; Ruiz et al., 2018). The neutron star has a radius on the order of 10 kilometres and rotates rapidly due to the conservation of angular momentum from its much larger radius progenitor (Özel & Freire, 2016). Therefore the period of the pulsar is very stable and a large external torque would be required to affect the pulsars spin. Pulsars are galactic sources and have extremely stable rotation periods making them excellent tools for probing the local interstellar medium. The beams of radiation are highly collimated and are only observed if they intersect with an observer’s line of sight. As such, radiation emitted by pulsars is observed in ‘pulses’ of radiation, which rotate anywhere between every 1 ms–5 s due to their rapid rotation speeds (Lorimer, 2008).

Fast radio bursts (FRBs) also produce radiation over short timescales (~ 1 ms) and are incredibly bright. Unlike pulsars however, FRBs have dispersion measures (see Subsubsection 1.3.2.2) much exceeding that expected from the Galaxy alone and therefore are an extra-galactic phenomenon. The mechanism behind FRBs is currently unknown, and there are two classes of FRBs: repeating; and non-repeating. The majority of FRBs do not repeat, and non-repeating FRBs are thought to be due to a cataclysmic event, whereas repeating FRBs are thought to be due to young magnetars (Margalit et al., 2018). However, it is not clear whether non-repeating FRBs have just not been observed long enough to detect repetition or if they are truly singular events. Localising FRBs could help in the determination of their origin but as most do not repeat determining their origin is difficult. The first FRB localised was the repeating FRB (121102) (Chatterjee et al., 2017). However, more recently the first non-repeating FRB (180924) was localised by Bannister et al. (2019). With more localisations in future the origin of these sources will become clearer. Due to their similarity in emission to pulsars, FRBs can also be used to probe the dispersion measures of our Galaxy, the intergalactic medium, as well as the ISM of their host galaxies and galaxy halos.

Pulsars and FRBs are both useful tools due to their ability to probe the intervening materials that their radiation travels through. This is because both emit large amounts of energy over short timescale (milliseconds - seconds) as seen in [Figure 1.4](#). These sources can be used to map the Galactic electron density, which allows models to reconstruct the plasma distribution of the ionized ISM. Models such as NE2001 and YMW16 ([Cordes & Lazio, 2002](#); [Yao et al., 2017](#)) have used pulsar dispersion measures along lines of sight to map the electron density of the ISM. With further FRB detections, these could also be included into newer models, increasing their reliability. Pulsars and FRBs also provide us with clear cut cases of extrinsic variability and their scintillation characteristics further elucidate the properties of the ISM.

1.2.2.3 Galaxies

The majority of sources found within radio surveys are extra-galactic in origin, specifically other galaxies. The radio emission observed from galaxies can be broadly grouped into two categories describing their origin; the radio emission produced via stellar processes within the galaxy disk (star-forming galaxies), and the radio emission produced via the accretion of matter onto the super-massive black hole at the galaxy centre (active galactic nuclei). Star-forming galaxies (SFG) are inherently not as bright as active galactic nuclei (AGN), due to the sources of their luminosity, as can be seen in the source counts found in [Figure 1.5a](#). The luminosity function, a measure of the number of sources as a function of brightness, can show us how these populations change as a function of flux density at a given frequency. [Figure 1.5b](#) shows that the number of SFG are greater at lower flux densities but quickly drop off and AGN become the dominant sources at high flux densities. SFG, have large emission sites, due to the fact that their emission comes from the whole galaxy (~ 30 kpc), whereas an AGN's emission will come from a compact core or site (~ 1 pc). Therefore the AGN is much more likely to be variable, this concept is discussed further in [Section 1.4](#). For the rest of this thesis we shall focus on AGN and their compact and extended components.



(a) Euclidean Normalised 1.4 GHz source counts.

(b) 150 MHz luminosity function.

Figure 1.5: (a) the euclidean normalised 1.4 GHz source counts from [Rawlings et al. \(2015\)](#), where the source counts are the number of sources observed at a given flux. Relevant colours and symbols as follows: total, blue squares; SF-powered galaxies, black stars; AGN-powered galaxies, red circles; AGN/SF hybrid, orange diamonds. (b) the 150 MHz luminosity function from [Sabater et al. \(2019\)](#). Relevant colours and symbols as follows: SF galaxies (blue diamonds); AGN (red circles); and the 1.4 GHz converted luminosity function (dashed lines).

1.2.3 Active Galactic Nuclei

A super-massive black hole (SMBH) is suspected to reside at the centre of every galaxy. The accretion of matter (the process by which material loses gravitational energy and collapses in onto the SMBH) liberates an enormous amount of energy from the galaxy nucleus, giving rise to the well known active galactic nuclei (AGN). The energy released from an AGN can be observed across various bands of the electromagnetic spectrum, however, AGN are particularly well known for their radio-loud activity. The structure of an AGN can be seen in [Figure 1.6](#), the SMBH is surrounded by an accretion disk which is fed by a torus of matter. The figure also shows two regions called the Broad Line Region (BLR) and the Narrow Line Region (NLR). The BLR is closer to the SMBH and the material is rotating around the black hole at high speeds ($\sim 10,000$ km/s) and the spectral lines are broadened by a Doppler shift from the rotation. NLRs are further away from the SMBH and do not rotate as quickly (~ 500 km/s), so the

matter does not experience the same magnitude of Doppler shift and so the spectral lines are narrower than those emitted from the BLR. Figure 1.6 also shows the two sub-groups of AGN: radio-quiet AGN (bottom half) and radio-loud AGN (top half), this classification is called the unification model of radio galaxies (Antonucci, 1993; Elitzur, 2012; Netzer, 2015).

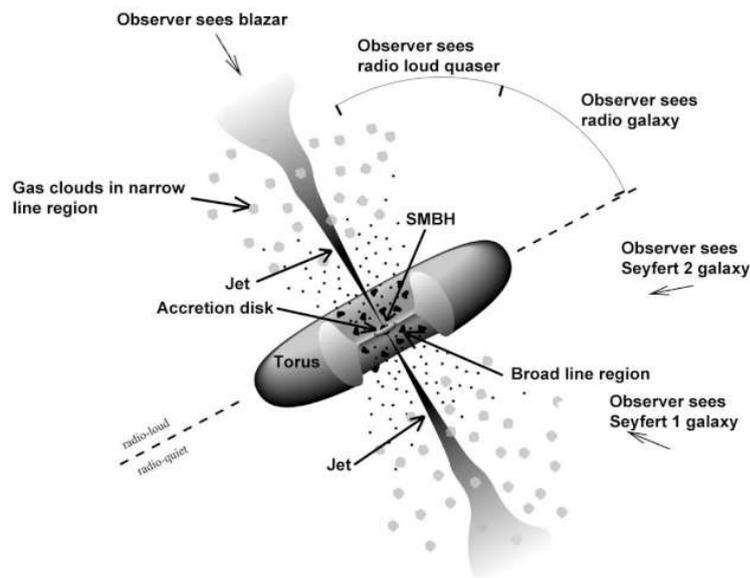


Figure 1.6: The *unification* model for radio galaxies, which divides the types of radio galaxies into radio- loud and quiet depending on the magnitude of their luminosities. Radio galaxies are then further broken down by the viewing angle of the observer. Also shown is the structure and emission sites of these galaxies. Source: FERMI³

Radio-loud AGN (Figure 1.7) produce a pair of jets which are relativistic, collimated, plasma outflows. The jets are launched from the galactic centre and break through the host galaxy out into the intergalactic medium (IGM) where they eventually terminate to form cocoons of shocked plasma, e.g. the radio lobes. A ‘radio core’, the jets, and the radio lobes all make up the defining features of a radio galaxy, which are observed and studied by radio facilities all over the world. Radio galaxies emit via synchrotron radiation, which is produced by electrons gyrating within a magnetic field. This forms a radio continuum that is *typically* well modelled by a power-law over a ~ 100 MHz – ~ 100 GHz frequency range.

³fermi.gsfc.nasa.gov/science/eteu/agn/

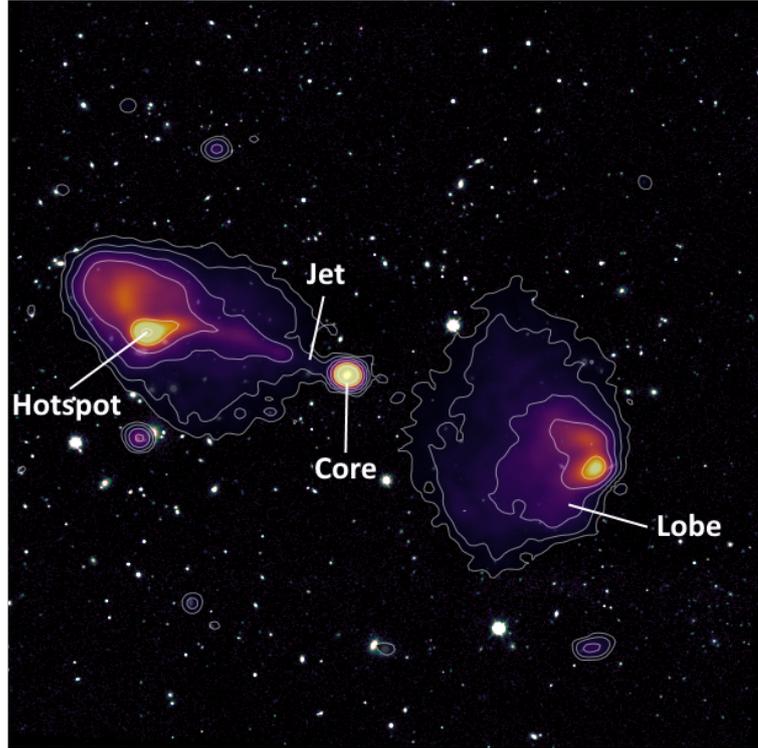


Figure 1.7: Radio galaxy PKS 2250-351 at a redshift of $z=0.2115$ observed using the Australian Square Kilometre Array Pathfinder (ASKAP) at a radio frequency of 887 MHz, the angular size of the source is 5.92 arc minutes. Labelled in this figure are the typical features of radio galaxies such as the core, jet, lobe and hotspot. Note that the counter-jet is not visible due to the orientation of the galaxy. More details on this source can be found at [Seymour et al. \(2020\)](#). Image compiled using code from Benjamin Quici.

AGN have been observed to exhibit both intrinsic and extrinsic variations, although distinguishing the divide has been historically difficult. Initially, it was assumed that the variations were intrinsic to the AGN, however, doubts of this statement first aired when it was found that some radio galaxies synchrotron emission was exceeding the brightness temperature limit. Radio sources emitting through synchrotron radiation have an upper limit on their brightness temperatures of $T_b = 10^{12} K$ ([Kellermann & Pauliny-Toth, 1969](#)) due to inverse-Compton losses, also known as the Compton limit. Brightness temperature quantifies the temperature a blackbody would need in order to produce the same intensity as the source. The Rayleigh-Jean ($h\nu \ll kT$) brightness temperature (T_b) is defined

by

$$T_b = \frac{I_\nu c^2}{2k_b \nu^2}, \quad (1.3)$$

where I_ν is the Intensity of the source, k_b is the Boltzmann constant, ν is the frequency and c is the speed of light. If the Compton limit was only being exceeded in statistically few sources, these could be explained by Doppler boosting, a mechanism that boosts the apparent brightness of a galaxy when the beam axis is close to the viewing angle. However, the number of sources that exceed this limit are so many that to be explained by Doppler boosting would suggest that jets are preferentially directed towards Earth. As this is an unlikely conclusion, astronomers looked for other solutions to how sources could be exceeding the Compton limit.

1.2.4 Doppler Boosting

Doppler boosting is a mechanism that causes an apparent change in intensity of some incident radiation, depending on the direction in which the emitting electron is travelling with respect to an observer's line of sight. The intensity of the beam is largest when the source of radiation (an electron) is travelling close to an observer's line of sight, and falls off as the angle (θ) increases between the line of sight and the beam axis. Doppler boosting becomes significant only at relativistic speeds and its effect is often observed in AGN jets. Some fraction of radio sources will have a jet directed towards an observer and a counter-jet directed away from an observer and this apparent difference in the jet velocities will produce an difference in the jet intensities. The jet and counter-jet are Doppler boosted and Doppler dimmed respectively. Whereas, in the case where the jets are close to perpendicular to an observer, they will exhibit little to no Doppler boosting, See [Figure 1.8](#) for a comparison.

The amount that a beam is affected by Doppler boosting is called the Doppler factor, \mathcal{D} , which is given by:

$$\mathcal{D} = \frac{1}{\gamma(1 - \beta \cos\theta)} \quad (1.4)$$

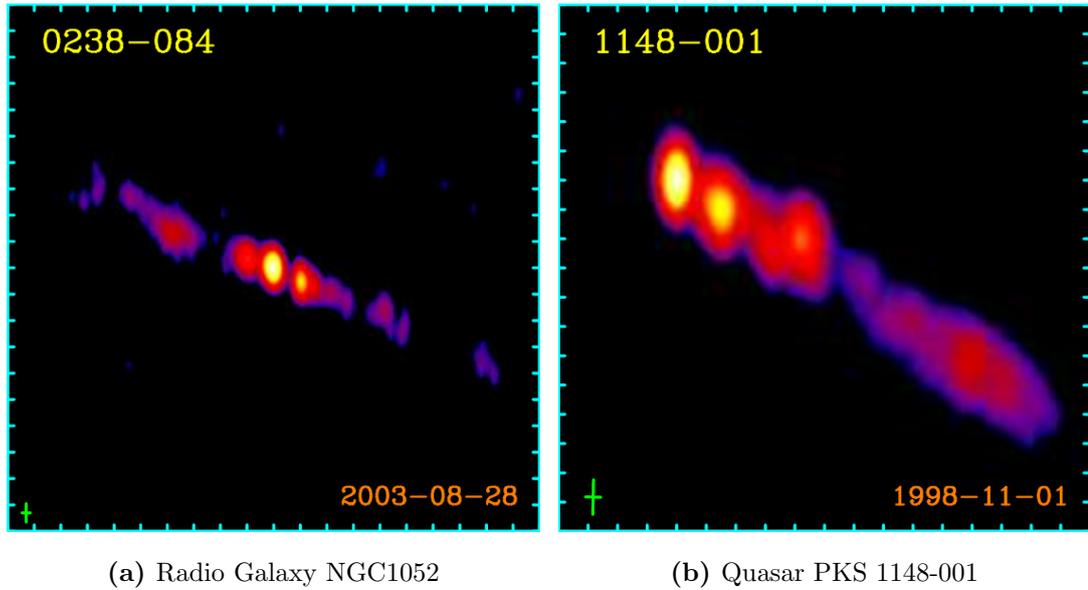


Figure 1.8: Comparison between (a) a non-Doppler boosted and (b) Doppler boosted source observed using 15 GHz images with VLBA from [Kellermann et al. \(2007\)](#). (a) The radio galaxy NGC1052 is almost perpendicular to our line of sight and so the jets have little to no Doppler boosting and look symmetric. The tick marks are two milli-arcseconds. (b) The Quasar PKS 1148-001 jets are Doppler boosted and the intensity of the observed radiation will be brighter and dimmer for the jet and counter jet respectively. The counter jet is not visible in this image. The tick marks are one milli-arcsecond.

where $\beta = v/c$, γ is the Lorentz factor given by $(1 - \beta^2)^{-1/2}$ and θ is the angle between the observer and the jet. If we infer the size of the source from the detected unresolved component then the observed brightness temperature (T_b) for a source (T'_b) that is being Doppler boosted is given by:

$$T_b = T'_b \mathcal{D} \quad (1.5)$$

However, if we assume that a source is intrinsically varying we can infer the linear source size (R) by invoking the light travel time argument: $R = c\Delta t'$ for the sources rest frame, which results in a brightness temperature derived from the intrinsic variations (T_{var}) with an extra \mathcal{D}^2 factor:

$$T_{var} = T'_b \mathcal{D}^3 \quad (1.6)$$

From this we can see that Doppler boosting has a significant effect on the brightness temperature and from [Equation 1.4](#) we can see that the Doppler fac-

tor is a maximum when the jet is pointed directly along our line of sight ($\theta = 0$). Using VLBI observations [Valtaoja et al. \(2003\)](#) were able to determine that this Doppler boosting could account for the large brightness temperature of up to $10^{13.5} K$ for quasars, $10^{11.5} K$ for radio galaxies, and up to a maximum of $10^{15} K$ for intrinsically varying sources. [Hovatta et al. \(2009\)](#) found a maximum variability Doppler factor of $\mathcal{D} = 35.5$, which could also account for brightness temperatures up to $\sim 10^{16} K$. This result accounted for moderately varying sources. However, variations on shorter timescales (hours to days) imply much larger brightness temperatures ($10^{18} K$ to $10^{21} K$) which cannot be reconciled with the assumptions of intrinsic variability and Doppler boosting.

1.2.5 Intra-day Variability

Intra-day variability (IDV) was first discovered by [Heeschen et al. \(1987\)](#) in a survey of radio variability in extra-galactic sources. In this survey some flat spectrum radio sources were found exhibiting significant flux variations on time scales of days or less. Using intrinsic light travel time causality arguments ($R < c\Delta t$), IDV constrains the source size for these variations to micro-arcseconds. The small source size imposed by intrinsic variations implied a large brightness temperature for these sources ($10^{18} K$ to $10^{21} K$). Many sources have been found that cannot be totally reconciled by restrictions imposed by intrinsic variations. [Kedziora-Chudczer et al. \(1997\)](#) found that if the source PKS 0405-385 was only intrinsically varying then the source would have a brightness temperature of $\sim 10^{21} K$. For Doppler boosting to explain this brightness temperature, a Doppler factor of up to 10^3 would be required ([Kedziora-Chudczer et al., 2001](#)), whereas typical Doppler factors range from 3-10 ([Qian et al., 1991](#)) for superluminal sources. Similarly [Hovatta et al. \(2009\)](#) found Doppler factors ranging from 0.3-35.5 using a combination of observations at 22-37 GHz and VLBI data. [Rickett et al. \(1995\)](#) also found that if source 0917+624 was intrinsically varying it would have a large brightness temperature of $10^{18} K$ and be so small ($0.1 \mu\text{as}$) that we would expect to see Diffractive InterStellar Scintillation (DISS, see [Subsubsection 1.4.2.1](#)) at

20cm, which was not observed. Instead they demonstrate that if Refractive Inter-Stellar Scintillation (RISS, see [Subsubsection 1.4.2.2](#)) in the ISM were the cause of the IDV it would produce a brightness temperature of up to $\sim 10^{12}K$. Under the assumption that ISS is the cause of these variations, there is no longer as strict a source size limit and so the brightness temperature becomes much more reasonable ($10^{13}K$ to $10^{15}K$). [Rickett et al. \(1995\)](#) also make a fundamental point that if the intrinsic argument is suggesting such small source sizes, then one should also expect to see ISS, and that further studies that only assume intrinsic variations should also consider ISS within their analysis. At this point, it is clear that the variations and brightness temperatures cannot be solely reconciled by the Doppler boosting of intrinsic variations, and so it was proposed that extrinsic variations could instead be the cause of the variability.

Later [Jauncey et al. \(2000\)](#) found that PKS 0405-385's variability pattern arrived at different times (140 seconds delayed) at two telescopes separated by 10,000 kilometers. This time delay is too large to be light arrival time delay and instead was found to be due to the ISM's transverse velocity of approximately 75kms^{-1} . Further, [Jauncey & Macquart \(2001\)](#) and [Rickett et al. \(2001\)](#) observed the modulation of IDV source 0917+624 and found that its timescale was changing on annual cycles. This result was due to changes in the relative velocity between Earth and the scattering medium as the Earth orbited the Sun. Due to the discovery of the regular six month cycles in the variations, we now know that IDV is mainly due to the ISS and we can shift our focus away from intrinsic causes. If ISS is the root cause of this variability a Doppler factor of 10^3 would still be required to boost the emission from $10^{12}K$ up to $\sim 10^{15}K$ ([Equation 1.5](#)). Evidence found from several sources ([Dennett-Thorpe & de Bruyn, 2000](#); [Rickett et al., 2002](#); [Bignall et al., 2003](#)) suggest that there is a bias in AGN showing rapidly varying scintillation and that they pass through a much closer phase screen at ~ 30 pc. This would mean that for ISS the brightness temperatures would also be approximately 30 times smaller when compared to a typical phase screen distance of 1 kpc. For ISS to explain these large brightness temperatures a Doppler factor in the range of $\sim 10^2$ would be required, which is

much more commonly found ([Hovatta et al., 2009](#)).

It should now be clear that intrinsic variation could not reconcile the large brightness temperatures for IDV sources. Instead variability is dominated by ISS in combination with small Doppler factors with perhaps some small component pertaining to intrinsic variations of the source ([Rickett, 2002](#)). In the cases presented, the ISM and the effects of ISS have been shown to be both a useful tool and a hindrance for astronomers, due to these reasons it is in our best interest to fully understand the media that we are observing through to better understand extrinsic variability.

1.3 The Intervening Medium

When observing sources, there are several different media that intersect our line of sight: the ionosphere, interplanetary medium (IPM), interstellar medium (ISM) and the intergalactic medium (IGM). The ionosphere medium is the ionized part of Earth’s upper atmosphere at a distance ranging from 80-1000km and an average electron density of $\sim 10^5 \text{ cm}^{-3}$. The IPM is what fills in the gap between the planets and the Sun, it starts at the edge of the Earth’s atmosphere and ends at the Heliosphere. The IPM is composed of hot plasma ejected from the Sun, with the particle density of the solar wind ranging from $5 - 20 \text{ cm}^{-3}$. The ISM is the physical space within the Galaxy outside of our solar system and is made up of several warm, cold, neutral and ionised components, with varying densities, states and temperatures. The IGM is the medium that exists between galaxies, it is incredibly diffuse with the ‘warm-hot’ intergalactic medium having an estimated density of $\sim 1 \times 10^{-6} \text{ cm}^{-3}$. Each medium affects signals differently due to different typical length-scales of inhomogeneities within the medium, the distance to the medium and the frequency of the observation. These factors determine the critical angles and timescales of the variability.

The ability to use these media as tools has been demonstrated by [Macquart & Jauncey \(2002\)](#) who use the relative motion of the Earth and the scattering

medium to learn about the source’s structure. Anything smaller than the critical scattering angle (e.g. a compact source or component), will be variable, thus allowing the authors to investigate the structure of IDV sources down to micro-arcsecond resolutions. This also allows them to determine whether the observed emission is from the central compact core or from the base of the jets. Similarly, [Chhetri et al. \(2018a\)](#) used interplanetary scintillation (IPS) to identify and measure sub-arcsecond compact components of low frequency radio sources without the need for long baseline interferometry. This allowed the authors to determine statistics about the population, specifically they found that peak-spectrum radio sources were the dominant strongly scintillating (compact) population. Further examples of interstellar scintillation being used to probe extra-galactic sources can be found in [Dennett-Thorpe & de Bruyn \(2003\)](#); [Bignall et al. \(2003\)](#); [McCulloch et al. \(2005\)](#); [McCallum et al. \(2009\)](#). To be able to use the intervening media as a tool or to predict its effects on our sources, we have to be able to understand the structure of the intervening media and how it can be mapped.

The MWA typically creates images averaged over two minutes so that scintillation induced by variations in the ionosphere and interplanetary media occur on timescales (seconds to minutes) averaged over in the imaging process. On the other hand, the IGM causes fluctuations on much longer timescales (years to decades) than most reasonable observing timescales. Furthermore, due to the diffuse nature of the IGM, it does not contribute significantly to the variability observed in radio galaxies. This results in the ISM being the ‘Goldilocks’ medium that has the right lengthscale and induces variability which occurs on typical observing timescales.

1.3.1 The Interstellar Medium

The ISM is largely composed of hydrogen (90.8% by number), which ranges in temperatures and states: [Table 1.2 \(Ferrière, 2001\)](#). As a large percentage of electron interactions occur due the Warm Ionised Medium (WIM) we can use the WIM as a proxy for the distribution of electrons. The WIM can be traced using

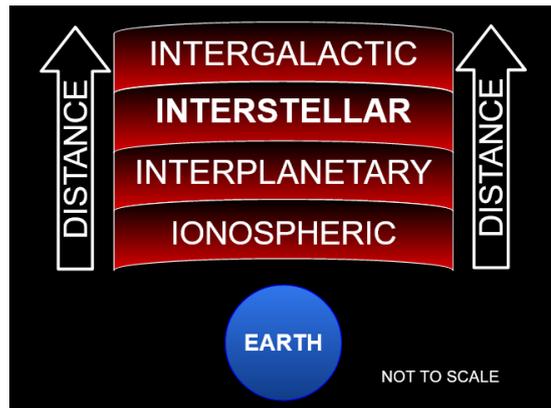


Figure 1.9: Some of the different media intersecting our line of sight, as a function of increasing distance. From closest (bottom) to furthest (top) the mediums are: ionospheric, interplanetary, interstellar, and intergalactic.

dispersion measures (DMs) of pulses passing through and its hydrogen alpha ($H\alpha$) emissions. The dispersion measure is the integrated column density of electrons along a line of sight, often measured using pulsars and FRBs. The strength of the optical emission from $H\alpha$ traces ionised regions of hydrogen, it can also be linked to the emission measure, which measures the square of the electron density integrated along a line of sight. Several models of our Galaxy have been created to try and recreate the distribution of the ISM to best predict variability. These models use maps of DMs, $H\alpha$ emission and emission/absorption from HII regions to recreate the electron content of the Galaxy. Models such as TC93, NE2001 and YMW16 (Taylor & Cordes, 1993; Cordes & Lazio, 2002; Yao et al., 2017) have used different complex Galaxy models, over time have increased the number of measurements included. However, these models have all been created with a focus on predicting DISS and DMs for the pulsar and FRB communities. Whereas for galaxies, which are larger in angular scale than the point-like pulsars, are rarely affected by DISS, and instead are typically found exhibiting RISS. Not much development has focused on RISS, as this scintillation has been seen to be much less apparent within observations due to its weak nature and long timescales (Rickett, 1986), but as technology improves we take larger and longer surveys over more epochs where RISS becomes increasingly apparent within data sets. To build a model for RISS we must first understand the effects that the ISM imparts onto propagating signals.

Components	Volume (%)	Temp (K)	Density (N/cm ³)	State	Detection
Cold Neutral (CNM)	1–5%	50–100	20–50	HI	21cm absorption
Warm Neutral (WNM)	10–20%	6000–10000	0.2–0.5	HI	21cm emission
Warm Ionised (WIM)	20–50%	8000	0.2–0.5	HII	H α & DM
H II regions	< 1%	8000	10 ² – 10 ⁴	HII	H α & DM
Hot Ionized (HIM)	30–70%	10 ⁶ –10 ⁷	10 ⁻⁴ –10 ⁻²	Ionized	X-ray & UV

Table 1.2: Components of the interstellar medium. The table shows several components of the ISM, their respective volumes, temperature, density, state and method of detection. These components are detected through several different means: H α optical emission lines, dispersion measures (DM), absorption or emission of the 21cm line, X-ray emission, and UV absorption. This table has been created using data from [Ferrière \(2001\)](#).

1.3.2 Mapping Turbulence and Inhomogeneities

In the previous sections we have shown how the intervening media can be both a hindrance and a tool for astronomers. To be able to predict how much scintillation and scattering we see, we must be able to map turbulence and inhomogeneities within the ISM. To do this we can use relations that describe the density fluctuations along a line of sight.

1.3.2.1 Density Fluctuation Relations

We can measure the density of electrons along a line of sight using the dispersion measure (DM). This measures the electron number density (n_e) between an observer and the source at distance D (usually a pulsar):

$$DM = \int_0^D n_e dS \quad \text{pc cm}^{-3} \quad (1.7)$$

Whilst DMs are usually obtained from pulsar measurements we can use H α , thermal radio emission and radio recombination lines to try and find the change

in density along a line of sight. The emission measure (EM) quantifies the electron density squared between an observer and a source at distance D :

$$\text{EM} = \int_0^D n_e^2 ds \quad \text{pc cm}^{-6} \quad (1.8)$$

The column density of electrons in the ISM can be easily mapped along pulsar line of sights. However, the variation in electron densities along a line of sight are not so easy to map. When we measure number density of electrons, we are assuming that they are evenly distributed along a line of sight. However, this may not be true and the density fluctuations (δn_e) along a line of sight are hard to predict: see [Cordes et al. \(1991\)](#); [Armstrong et al. \(1995\)](#); [Cordes & Lazio \(2002\)](#) for a full review. Therefore assumptions are made about the turbulence seen in the ISM. We can characterise the fluctuations by a power-law wavenumber spectrum:

$$P_{\delta n_e}(q) = C_n^2 q^{-\beta}, \quad \frac{2\pi}{l_0} \leq q \leq \frac{2\pi}{l_1}, \quad (1.9)$$

where q is the wavenumber, C_n^2 is the spectral coefficient and in this case indicates the strength of the fluctuations. β is a power-law index which describes how the wave is modulated between the outer and inner scales l_0 and l_1 respectively. Another common assumption is that when energy is transferred between smaller and larger scales the resulting spectrum is a Kolmogorov power law ($\beta = 11/3$), which approximates the ISM well in general terms ([Armstrong et al., 1995](#)). We can determine a scattering measure (SM) from radio scattering observations that allows us to obtain a relation to the distribution of the electron density fluctuations along a line of sight.

$$\text{SM} = \int_0^D C_n^2 ds \quad \text{kpc m}^{-20/3} \quad (1.10)$$

We can use the scattering measure to model the electron density fluctuations (δn_e), however, it cannot be measured directly and instead a proxy is required. The two proxy's that can be used are: DMs and hydrogen alpha.

1.3.2.2 Dispersion Measure

In pulsar and FRB studies, the dispersion measure (DM) is used as a measure of how heavily a signal is dispersed by the column density of electrons. Due to this effect we can also use pulsar and FRB signals to work backwards to measure

the DM along a line of sight. When radiation travels through a medium that has a different refractive index the radiation is refracted and its group velocity changes based upon the frequency of the radiation passing through. Specifically higher frequencies are slowed down less than lower frequencies. The amount each pulse is delayed (Δt_d) can be calculated from the following equations (Lorimer & Kramer, 2005):

$$t_d = \frac{e^2}{2\pi m_e c} \times \frac{\int_0^d n_e dl}{f^2} = 4.15 \times 10^6 \text{ ms} \times \frac{DM}{f^2}, \quad (1.11)$$

where e is the charge of electron, m_e is the mass of an electron and c is the speed of light. To work out the difference in the arrival time of two pulses (Δt_d) we can calculate the difference between two frequencies:

$$\Delta t_d = 4.15 \times 10^6 \text{ ms} \times (f_1^{-2} - f_2^{-2}) \times DM \quad (1.12)$$

However, when a pulse travels through inhomogeneities, the signal is also scatter-broadened on timescales of τ_d (Equation 1.13). The broadening of the pulse can be seen in the exponential tail in Figure 1.10. This effect is also dependant on frequency and scales as $\tau_d \propto f^{-4} d^2$, where d is distance to the source, and f is the frequency.

$$\tau_d = \frac{\theta_{\text{scatt}}^2 D_{\text{eff}}}{2c} \quad (1.13)$$

Using observational data we can produce a functional relation between the pulse temporal broadening to DM:

$$\log \tau_d = a + b \log(DM) + c \log(DM)^2 - \alpha \log(\nu), \quad (1.14)$$

where ν is the frequency in GHz and τ_d is the pulse delay in milliseconds (ms). Bhat et al. (2004) obtained the following coefficients: $a = -6.59$, $b = 0.129$, $c = 1.02$ and $\alpha = 3.86 \pm 0.16$.

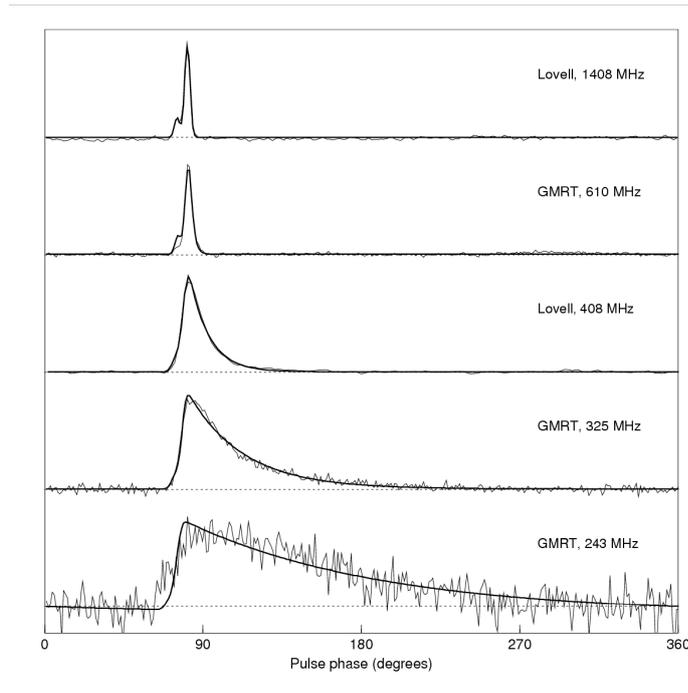


Figure 1.10: Pulse broadening over frequency. The diagram shows pulsar B1831-03's pulse profile at five different frequencies. The diagram emphasises the effect of frequency on the scattering and shows that at lower frequencies it has a significant effect on the pulse profile (also shown in Equation 6). The solid line is the exponential model fit for the data. [Lorimer & Kramer \(2005\)](#)

1.3.2.3 Hydrogen Alpha

Hydrogen alpha ($H\alpha$) is a Balmer series optical line emitted when an electron transits from the 3rd to 2nd energy level of a hydrogen atom. However, the probability of the atom being ionised is much higher than that of being excited to the 3rd level. Due to this, $H\alpha$ is usually produced when the hydrogen atom is reformed; the electron can then cascade down, producing $H\alpha$ emission. This is why $H\alpha$ is a tracer of ionised regions of hydrogen. These ionised regions of hydrogen contribute to the WIM and therefore $H\alpha$ is a proxy for electron density, which can be linked to the SM. We use [Haffner et al. \(1998\)](#), (eq. 1) and [Cordes & Lazio \(2002\)](#) (eq. 16) to convert between $H\alpha$ intensity ($I_{H\alpha}$) and SM:

$$\text{SM} = \left(\frac{I_{H\alpha}}{198R} \right) T_4^{0.9} \frac{\varepsilon^2}{(1 + \varepsilon^2)} \ell_0^{-2/3} \text{ kpc.m}^{-20/3} \quad (1.15)$$

Where $I_{H\alpha}$ is the intensity of the $H\alpha$ measured in Rayleighs, T_4 is the gas temperature in units of 10^4 K, ε is the fractional variance of n_e inside clouds

($\epsilon^2 = \langle (\delta n_e^2) \rangle / \bar{n}_e^2$) and ℓ_0 is the outer scale of turbulence in units of pc. The H α map created by [Finkbeiner \(2003\)](#) can be seen in [Figure 1.11](#). This scattering

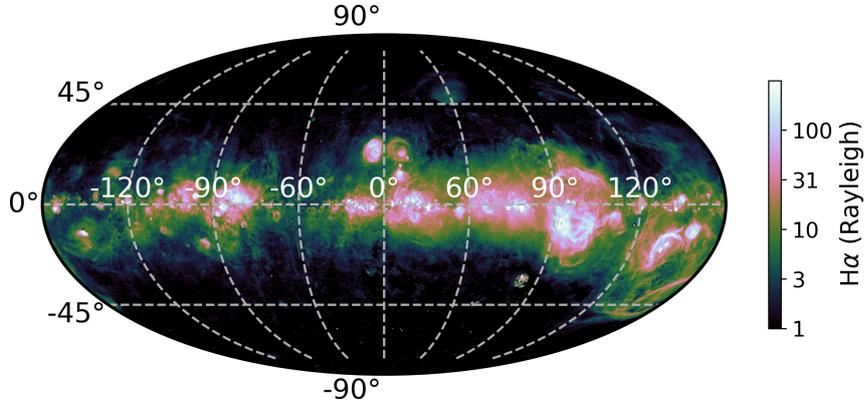


Figure 1.11: Hydrogen alpha map created by [Finkbeiner \(2003\)](#), measured in Rayleighs, displayed in galactic coordinates.

measure can be used to approximate the diffractive lengthscale (r_{diff}) as described further in [Macquart & Koay \(2013\)](#); Equation 7a:

$$r_{\text{diff}} = 3.7 \times 10^9 \left(\frac{\lambda}{1m} \right)^{-6/5} \left(\frac{\text{SM}}{10^{12} \text{m}^{-17/3}} \right)^{-3/5} \quad (\text{m}), \quad (1.16)$$

where λ is the wavelength and SM is the scattering measure. The diffractive lengthscale is one of the scattering lengthscales that can be used to determine the strength of variability an observer would see.

1.4 Scintillation and Scattering

Understanding the effects of scintillation and scattering is relevant now more than ever with the construction of low frequency radio telescopes such as the Murchison Widefield Array (MWA; [Tingay et al., 2013](#)), the Long Wavelength Array (LWA), the Karoo Array Telescope (MeerKAT; [Jonas & MeerKAT Team, 2016](#)), the LOW-Frequency ARray (LOFAR; [van Haarlem et al., 2013](#)), the Australian Square Kilometre Array (SKA) Pathfinder (ASKAP; [Johnston et al., 2007](#)) and, in future the (SKA; [Dewdney et al., 2009](#)). The frequencies that these telescopes observe

at ($\leq 10\text{GHz}$) are firmly within the strong regime of scattering as can be seen in Figure 1.12.

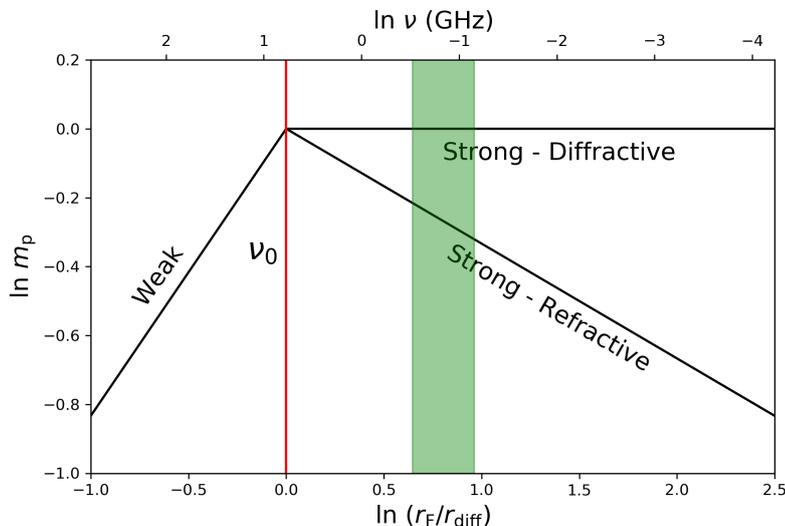


Figure 1.12: This figure has been recreated from Narayan et al. (1992), displays the regimes of weak and strong (refractive and diffractive) scattering. It shows the change in modulation as a function of the lengthscales r_F and r_{diff} as well as the frequency. The green band shows the frequency range for the MWA (70-300 MHz) and the red line shows the transition frequency of approximately 5 GHz for this line of sight.

Fluctuations in the refractive index within an intervening medium causes scintillation and scattering effects and the strength of these fluctuations can be used to define two different regimes of scattering: weak and strong. As their name implies they also impart different strengths of modulation onto the observed flux density, further, each regime occurs on different time and length scales (Narayan et al. (1992), Table 1). Observing extra-galactic radio sources means that we will be observing through the ISM, which imparts an effect onto all signals observed. The fundamentals of scintillation are well established (Rickett, 1990; Narayan et al., 1992; Walker, 1998) and there is much literature on different models of the ISM which are used to predict the variability we expect to observe when looking through the ISM (Taylor & Cordes, 1993; Cordes & Lazio, 2002; Yao et al., 2017). In this section we will summarise the theory of scintillation and scattering.

The ISM is the matter between stars within our Galaxy. It is turbulent and contains many inhomogeneities and when radiation propagates through these

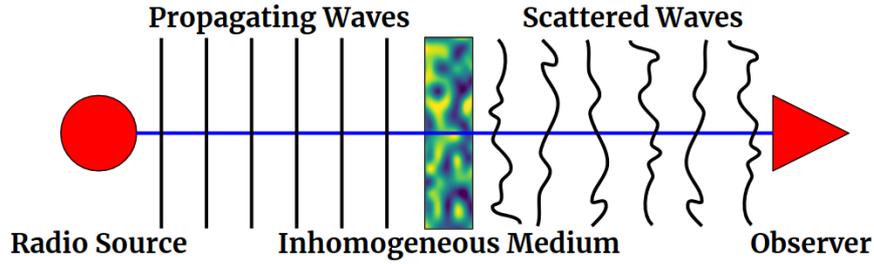


Figure 1.13: Thin screen model. Figure shows the effect of an inhomogeneous medium collapsed into a thin screen has on a wave when it passes through it.

inhomogeneities, scintillation and scattering effects are observed. These effects are dependent upon the the electron content (n_e) between the source and observer, the observing frequency (ν), the size of the source (θ_S) and the relative transverse velocity (v) of the ISM. A full three-dimensional treatment of the physical optics involved would require more parameters than are practically able to be measured along any line of sight. A more tractable solution is to assume that all the scattering occurs at a single screen along the line of sight - the so-called thin screen model (Scheuer, 1968). The thin screen lies some distance (D) between the target and observer (Figure 1.13). This model can be used to predict what happens to passing radiation and the observed wavefield is described by the Fresnel-Kirchoff integral:

$$\psi(X, Y) = \frac{e^{-i\pi/2}}{2\pi r_F^2} \iint \exp \left[i\phi(x, y) + i \frac{(x - X)^2 + (y - Y)^2}{2r_F^2} \right] dx dy, \quad (1.17)$$

where (X, Y) is the position on the observer plane, (x, y) is the position on the scattering screen plane, $\psi(X, Y)$ is the complex wave amplitude, and $\phi(x, y)$ is the phase change introduced by the screen. We can just focus on the two terms within the exponential; the first term in the exponential describes the phase change imparted by the medium, which is described by the diffractive lengthscale (r_{diff}). The second term is the phase change due to path length differences on the Fresnel scale (r_F , Equation 1.18),

$$r_F = \sqrt{\lambda D / 2\pi}, \quad (1.18)$$

where λ is the wavelength of the radiation and D is the distance to the scattering screen. When there are no significant phase changes ($\psi(x, y) = 0$) imparted by the screen r_F is the dominant scale with Fresnel angle - θ_F (Equation 1.19). Physically,

this happens when radiation is emitted from the source and at some point is scattered to the observer. The region where the scattered signal contributes coherently due to small path length change is called the first Fresnel zone. Signals from the first Fresnel zone are received from the Fresnel angle as defined by:

$$\theta_F = r_F/D, \quad (1.19)$$

where r_F is the Fresnel scale and D is the distance to the phase screen. Outside of the first Fresnel zone the phase oscillates rapidly with declining amplitude. These effects are purely due to the geometric time delays from signals travelling further distances and arriving at different times. This is represented in [Figure 1.14](#).

Alternatively, when $\psi \neq 0$ the medium imparts significant random phase fluctuations onto the wavefront. This can be described by the phase structure function ([Narayan et al., 1992](#)). The diffractive length scale (r_{diff}) is the transverse separation for which the root mean square (rms) phase difference is equal to one radian. We now have two scales r_{diff} and r_F with which we describe two different modes of scintillation. These scales are now used to define two regimes of scattering: weak and strong, where the name describes the modulations of a source as seen by an observer.

The two regimes are: weak scattering, where $r_{\text{diff}} \gg r_F$ the random phase fluctuations across the first Fresnel zone are small and r_F is still the dominant scale with only weak phase changes being imparted on the wavefront; and strong scattering, where $r_{\text{diff}} \ll r_F$, the random phase fluctuations vary on scales much smaller than the Fresnel scale (hence r_{diff} is the dominant scale) where the phase is approximately constant over regions on the phase screen with a radius of r_{diff} .

It is important to note that the limiting cases of weak and strong scattering simplify the mathematics needed to understand scintillation in these regimes. The transitional case is very tricky and is often avoided even though observations can be within this intermediate regime.

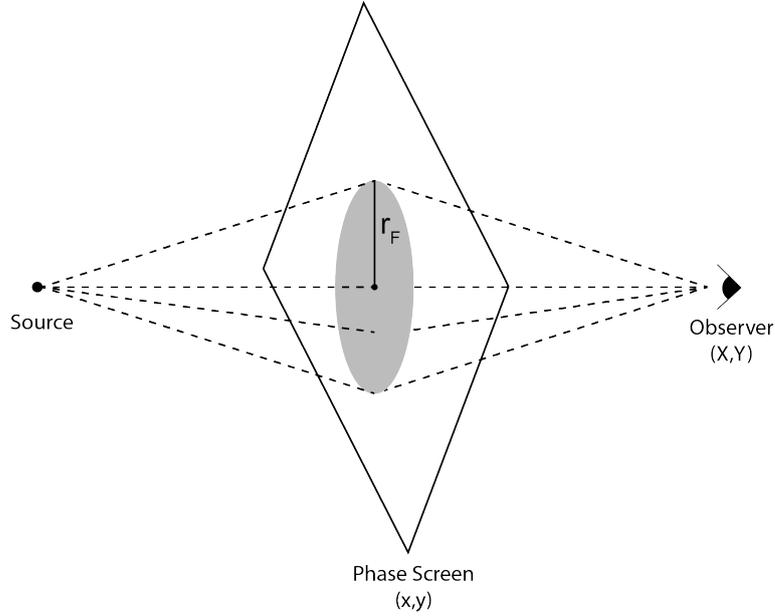


Figure 1.14: Representation of the first Fresnel zone on the scattering screen. The Fresnel scale describes effects seen from purely geometric time delay. A source’s signal can be received from any position on the plane. Within the first Fresnel zone of size r_F these signals add coherently. Diagram created by Jessica Wu (2019).

1.4.1 Weak Scattering

In weak scattering small variations of the flux occurs due to geometric time delays, the fluctuations occur on length scales of approximately r_F on the scattering screen. A point source is defined as having an angular size (θ_S) smaller than the critical scale of scattering θ_F ($\theta_S < \theta_F$). A modulation index (m) can be defined as the root mean square amplitude (σ) of the flux scintillation divided by its mean (μ).

$$m = \frac{\sigma}{\mu} \quad (1.20)$$

The point source modulation index (m_p) for weak scattering is calculated as,

$$m_p = \left(\frac{r_F}{r_{\text{diff}}} \right)^{5/6}, \quad (1.21)$$

where the timescale (t_F) for this scattering, assuming a transverse velocity (v) is:

$$t_F = \frac{r_F}{v}. \quad (1.22)$$

For an extended source ($\theta_S > \theta_F$) we can think of this as being composed of multiple “point sources”, shifted by some distance from the centre of the source. Due

to the random offset of each wave being emitted by multiple “point sources” the total amplitude received is suppressed by the size of the source (Equation 1.23).

$$m_e = \left(\frac{r_F}{r_{\text{diff}}} \right)^{5/6} \left(\frac{\theta_F}{\theta_S} \right)^{7/6} \quad (1.23)$$

From this we can compile scattering parameters for the weak scattering regime (Table 1.3). Weak scattering within the ISM occurs at $\nu \geq 5\text{GHz}$ which is larger than the frequencies that we are observing at $\nu \leq 3\text{GHz}$, so while weak scattering is important we will not further consider its effects.

Weak Scintillation	
Compact ($\theta_S < \theta_F$)	Extended ($\theta_S > \theta_F$)
$r_{\text{scint}} \approx r_F$	$r_{\text{scint}} \approx r_F(\theta_S/\theta_F)$
$t_{\text{scint}} \approx t_F$	$t_{\text{scint}} \approx t_F(\theta_S/\theta_F)$
$m_p \approx (r_F/r_{\text{diff}})^{5/6}$	$m_e \approx (r_F/r_{\text{diff}})^{5/6}(\theta_F/\theta_S)^{7/6}$

Table 1.3: Weak scattering parameters. Table showing the different parameters for a compact or extended source undergoing weak scattering.

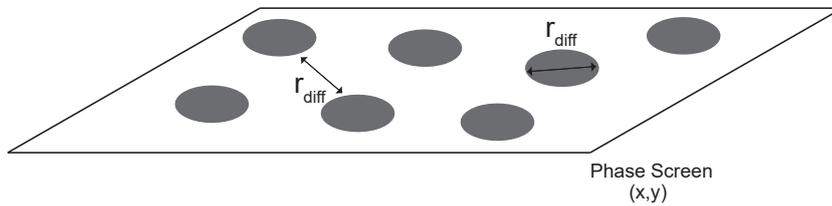


Figure 1.15: Representation of diffractive lengthscale, on the scattering screen. r_{diff} represents zones where the phase change due to inhomogeneities is less than one radian and is therefore coherent. These patches have approximate size r_{diff} as well as being spaced on average by r_{diff} . Diagram created by Jessica Wu (2019).

1.4.2 Strong Scattering

Strong scattering is characterised by phase fluctuations where $r_{\text{diff}} \ll r_{\text{F}}$ across the fresnel scale. There are two characteristic length scales: (i) r_{diff} and (ii) a larger lengthscale called the refractive lengthscale ($r_{\text{ref}}=r_{\text{F}}^2/r_{\text{diff}}$). For typical radio observations of decimetre wavelength and longer that observe the ISM; the dominant regime is strong scattering (Rickett, 1990). The variations due to strong scattering can be split into two separate regimes, diffractive and refractive scattering. Intensity fluctuations occur due to diffractive interstellar scintillation (DISS) on short timescales with large flux modulations, and refractive interstellar scintillation (RISS) on long timescales (days to decades) with smaller flux modulations.

1.4.2.1 Diffractive Scintillation

In diffractive scattering each point on the observer's plane receives radiation from a large number of points from the scattering, leading to multipath propagation. This is due to having coherent patches of size r_{diff} on the scattering screen (Figure 1.15), with each patch scattering radiation into a cone with angle θ_{scatt}

$$\theta_{\text{scatt}} = \frac{\lambda}{2\pi r_{\text{diff}}}, \quad (1.24)$$

as seen in Figure 1.16. An observer sees radiation from all patches within r_{ref} so that the image viewed is scatter-broadened on the scattering screen by angle θ_{scatt} (such that $r_{\text{ref}}=D\theta_{\text{scatt}}$) as defined in Equation 1.24.

Interference between the wavefronts from each r_{diff} size coherent patch gives rise to the intensity fluctuations seen by the observer. This change in flux occurs on very short length and timescales in comparison to refractive scintillation. However, the modulation is much stronger ($m_{\text{p}} \sim 1$). As r_{F} is no longer the dominant scale it no longer defines the critical angular scale for which a source is determined to be compact. Instead this is determined by the diffractive angular scale (θ_{diff}):

$$\theta_{\text{diff}} = r_{\text{diff}}/D. \quad (1.25)$$

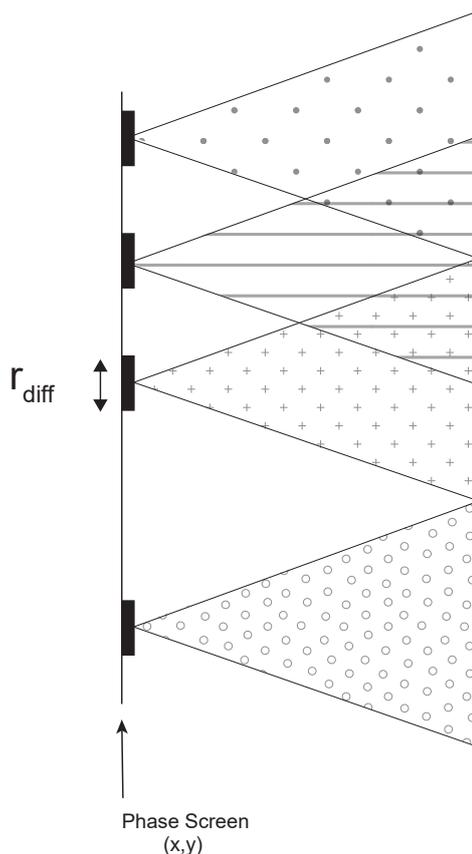


Figure 1.16: Coherent patches of r_{diff} act like slits and re-emit cones of light into angle θ_{scatt} . An observer sees patches within r_{ref} . Diagram created by Jessica Wu (2019).

The corresponding fluctuation timescale is:

$$t_{\text{diff}} = r_{\text{diff}}/v. \quad (1.26)$$

Similar to weak scattering, extended source in this case ($\theta_s > \theta_{\text{diff}}$ for diffractive scintillation) increases the timescale of scintillation and decreases the modulation index as outlined in Table 1.4. Signals from sources that are sufficiently compact (such as pulsars and FRBs) are strongly affected by DISS.

1.4.2.2 Refractive Scintillation

Refractive scintillation is caused by large scale inhomogeneities across the lengthscale r_{ref} on the scattering screen which act as a lens that focuses or defocuses the cones of radiation produced by the r_{diff} sized coherent patches (Figure 1.17). When the lens acts as a focus, cones from outside of r_{ref} are focused into our line of sight and an observer would receive more flux as a result. The

Diffractive Scintillation	
Compact ($\theta_S < \theta_{\text{diff}}$)	Extended ($\theta_S > \theta_{\text{diff}}$)
$r_{\text{scint}} \approx r_{\text{diff}}$	$r_{\text{scint}} \approx r_{\text{diff}}(\theta_S/\theta_{\text{diff}})$
$t_{\text{scint}} \approx t_{\text{diff}}$	$t_{\text{scint}} \approx t_{\text{diff}}(\theta_S/\theta_{\text{diff}})$
$m_p \approx 1$	$m_e \approx \theta_{\text{diff}}/\theta_S$

Table 1.4: Diffractive scintillation parameters. This shows the different parameters for a compact or extended source undergoing diffractive scintillation.

opposite is also true for defocusing, as an observer would receive less flux as the cones are ‘focused’ out of their line of sight. Refractive scintillation occurs on much longer timescales than diffractive scintillation:

$$t_{\text{ref}} = r_{\text{ref}}/v. \quad (1.27)$$

Refractive scintillation also occurs on much longer length scale and also has a much weaker effect on the modulation:

$$m_p = (r_{\text{diff}}/r_F)^{1/3} \quad (1.28)$$

For an extended source we see the same effects as in diffractive scintillation of an extended source: a increase in the timescale and a decrease in modulation index. However, for refractive scintillation, the source size limit is much less strict with the source only having to be smaller than the scattering disk (θ_{scatt}) rather than the diffractive angular scale (Equation 1.25). The comparison of scattering parameters for a point and extended source are described in Table 1.5.

1.4.3 Application

From this point we have two options with what to do with this information about scintillation. The first is that we could measure variable objects and work backwards to investigate the intervening material between us, the observer, and the object. The second is to use models/assumptions of the turbulence and distribution of material in the line of sight of the object to predict variability. The first option requires a lot of time and line of sight measurements, whilst

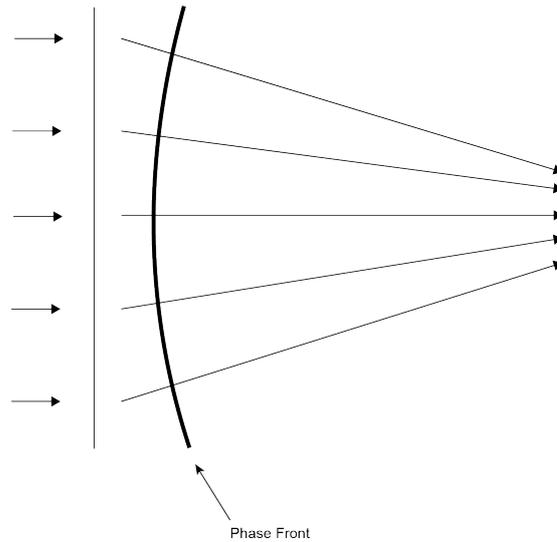


Figure 1.17: Cones of radiation emitted by the diffractive patches can be focused/defocused by a large scale phase front, occurring on scale r_{ref} . Where an observer would view more/less patches depending on whether the phase front is being focused towards or defocused away from an observer. The above diagram shows an example of diffractive patches being focused by a lens of size r_{ref} . Diagram created by Jessica Wu (2019).

the second can be done using easily obtainable models. In this dissertation we have chosen to pursue the second option. As previously mentioned models such as TC93, NE2001 can produce scintillation parameters and YMW16 maps DMs and electron density. However, these models are all focused on the short time scale DISS regime. While the outputs of these models could be used for other types of scintillation, this would not be a simple task and would require significant time investment from a novice user. As there is currently no model specifically focused on the longer timescale RISS, the goal of this project is to fill this gap with a model for this regime, as further discussed in the [Section 1.5](#). To complete this goal we first have to be able to relate the turbulence and Inhomogeneities in the ISM to the amount of scattering seen, this can be done through density fluctuation relations, discussed further in [Subsection 1.3.2](#).

Refractive Scintillation	
Compact ($\theta_S < \theta_{\text{scatt}}$)	Extended ($\theta_S > \theta_{\text{scatt}}$)
$r_{\text{scint}} \approx r_{\text{ref}}$	$r_{\text{scint}} \approx r_{\text{ref}}(\theta_S/\theta_{\text{scatt}})$
$t_{\text{scint}} \approx t_{\text{ref}}$	$t_{\text{scint}} \approx t_{\text{ref}}(\theta_S/\theta_{\text{scatt}})$
$m_p \approx (r_{\text{diff}}/r_F)^{1/3}$	$m_e \approx (r_{\text{diff}}/r_F)^{1/3}(\theta_{\text{scatt}}/\theta_S)^{7/6}$

Table 1.5: Refractive scintillation parameters. The different parameters for a compact or extended source undergoing refractive scintillation are listed.

1.5 Summary

Using telescopes to survey the radio sky has allowed astronomers to gain a greater understanding of our Universe; one piece of this puzzle is the variability found in radio galaxies. Variability is an important astrophysical tool that encodes information about a source, but we also must have the ability to decode the information. To do this we must understand more about the source and cause of the variability. From the literature we have found that the most extreme cases of variability (IDV) were due to extrinsic variability (RISS) from the ISM (Jauncey et al., 2000; Jauncey & Macquart, 2001; Rickett et al., 2001; Dennett-Thorpe & de Bruyn, 2003; Bignall et al., 2003; McCulloch et al., 2005; McCallum et al., 2009). Intervening media can also be used as a tool to investigate sources, as demonstrated by Macquart & Jauncey (2002) and Chhetri et al. (2018a). At sub-gigahertz radio frequencies the ISM is the dominant medium which can be mapped using dispersion measures and hydrogen alpha measurements.

The basis of this thesis is the combination of understanding sources and understanding the ISM. Using measurements and theory about the distribution of matter within the ISM, as well as theory about sources and how they scintillate, we can predict what we should observe. If this matches with the observational data, then the models are sufficiently accurate. If there are significant discrepancies then this informs us that the models are not sufficiently detailed or accurate to replicate what an observer would see.

Dr. Paul Hancock created a base model (SM2017) for predicting RISS using hydrogen alpha ($H\alpha$) as a proxy for scattering measure and then linking this to scattering theory. We will update this model RISS19. Hancock et al. (2019) also used SM2017 to demonstrate the effects of RISS on MWA radio observations of several regions. Within this paper it is found that indeed the $H\alpha$ is correlated with the modulation indices of point sources. We will then further develop the RISS19 model to predict the scintillation modulation index for an extended source size, as well as other upgrades detailed in Chapter 2. Further, a simulation will be created to take this RISS19 model and use it as a base to create population statistics about a given region of sky. We can then compare this to actual observational results to determine whether the RISS model and the underlying theory can accurately predict what is observed.

In Chapter 2 we will describe the base model that has been constructed by Dr. Paul Hancock and the improvements we have made to this model. In Chapter 3 we will show how a simulation was developed and the results it predicted. Finally in Chapter 4 we will summarise the results that we have found and what further improvements can be made.

Chapter 2

Scintillation and Scattering Models

2.1 Introduction

It has been demonstrated in the previous chapter that the effects that the ISM imparts can be combined with tracers of the ISM to create a model to predict the variability of scintillating radio sources. One such model has been created by Dr. Paul Hancock: the SM2017 model predicts the scintillation induced variability of radio sources from scattering theory based on their sky position and a map of the distribution of $H\alpha$. Its successor RISS19 has been updated from SM2017 to include new calculations for: Galactic distance, extended sources, and the scattering disk size. The purpose of this work is to considerably extend the capabilities of RISS19 based on the results from [Hancock et al. \(2019\)](#). We create a new branched version, TAU19, that removes dependency on the distance to the scattering screen by using the ‘*Bhat Relation*’ ([Bhat et al., 2004](#)) - a relation between dispersion measure and pulse temporal broadening. In this section we will discuss the basic program design, what it requires as an input and what it can output. We will also discuss the dependencies of each model, what we have done to limit these and what improvements have been made to upgrade each iteration of the model. All work discussed in [Section 2.2](#) has been completed by Dr. Paul Hancock, and I provide a brief review and description for the reader.

2.2 SM2017

The SM2017 model was written by Dr. Paul Hancock and was modified by me, as described in later sections ([Section 2.3](#)). Here we will concisely lay out how the SM2017 program works and what modifications were made to it. The flow chart for SM2017 can be seen in [Figure 2.1a](#). In summary SM2017, works by taking in positions and returning H α Intensity value for each specified position, which is then used to calculate different scattering parameters which are used to calculate the interstellar scintillation parameters. The program also has several default parameters and values shown in [Table 2.1](#).

Name	Value	Unit	Modifiable?
Frequency	185	MHz	Y
Screen Distance	1	kpc	N
Observed Time	1	years	Y
ISM Velocity	10^4	ms^{-1}	N

Table 2.1: SM2017 Defaults and Constants.

The model first takes Galactic coordinate positions from the user. This is input into the program with the hydrogen alpha map, the observing frequency and the observation time period. The program returns a corresponding H α intensity value and calculates the Fresnel Scale from the constants assumed. The H α intensity is then linked to the scattering measure, as described in [Subsubsection 1.3.2.3](#) and [Equation 1.15](#).

$$\text{SM} = \left(\frac{I_{\text{H}\alpha}}{198R} \right) T_4^{0.9} \frac{\varepsilon^2}{(1 + \varepsilon^2)} \ell_0^{-2/3} \text{ kpc.m}^{-20/3}$$

The SM2017 model uses the following values $\varepsilon = 1$, $T_4 = 0.8$ ([Haffner et al., 1998](#)), and $\ell_0 = 10^{18}\text{m} = 32\text{pc}$ ([Armstrong et al., 1995](#)). Using these values we obtain a simplified form of the linear relation between the scattering measure and the H α intensity in units of Rayleighs:

$$\text{SM} = 1.26 \times 10^{16} I_{\text{H}\alpha} \text{ m}^{-17/3}. \quad (2.1)$$

SM2017 then derives the diffractive scale r_{diff} through Kolomogorov modelling of the turbulence and the assumption that the scattering disk is within our Galaxy, as described in [Macquart & Koay \(2013\)](#) and [Subsubsection 1.3.2.3](#):

$$r_{\text{diff}} = 3.7 \times 10^9 \left(\frac{\lambda}{1\text{m}} \right)^{-6/5} \left(\frac{\text{SM}}{10^{12}\text{m}^{-17/3}} \right)^{-3/5} \text{ m.}$$

SM2017 then uses the two calculated scattering scales, r_{diff} and r_{F} , to calculate the scattering parameter ξ :

$$\xi = r_{\text{F}}/r_{\text{diff}}, \quad (2.2)$$

where r_{F} is the Fresnel scale assuming the screen is at a distance of 1 kpc ([Equation 1.18](#)). From this the point source modulation index can be calculated for refractive scintillation:

$$m_p = \xi^{-1/3} \quad \text{for } \xi > 1. \quad (2.3)$$

SM2017 also calculates two other parameters, the refractive timescale (t_{ref}) of the variation and the expected RMS variations over an observed time period. The refractive timescale is calculated by:

$$t_{\text{ref}} = \xi \left(\frac{r_{\text{F}}}{v} \right). \quad (2.4)$$

The refractive timescale is then used to calculate the RMS modulation variations (m_{rms}) for any given observed time period in years (t_{obs}):

$$m_{\text{rms}} = m_p \left(\frac{t_{\text{obs}}}{t_{\text{ref}}} \right) \quad \text{for } t_{\text{obs}} \leq t_{\text{ref}} \quad (2.5)$$

All of the calculated quantities, except the Fresnel and diffractive scales, can be output to a table for further analysis. This base model can return modulation results for singular or multiple positions. However, the model makes some basic assumptions about the scattering screen distance and there are several improvements that can be made to it.

2.3 RISS19

RISS19 is the published upgraded version of SM2017 A summation of the RISS19 program in visual form can be seen in [Figure 2.1b](#) and the default parameters and constants used in [Table 2.2](#). The work completed in this section was completed

Name	Default Value	Unit	Modifiable?
Frequency	185	MHz	Y
Screen Distance	N/A	kpc	Y
Observed Time	1	years	Y
ISM Velocity	10^3	ms^{-1}	Y

Table 2.2: RISS19 Defaults and Constants.

by me and builds upon work done by Dr. Paul Hancock, unless otherwise explicitly stated. The first improvement is to include error analysis within the code, as we want to be able to determine the significance of the results. To include errors, we used the associated error map of the hydrogen alpha map, and did basic uncertainty propagation to include errors in final calculated results.

One major difference between the two programs is the function to calculate a dynamic scattering screen distance based on position in the sky, rather than the static 1 kpc distance used in SM2017. As the Earth is not located at the centre of the Galaxy the amount of ISM one will be observing through changes as a function of viewing angle. Therefore, the scattering screen distance should be change with the viewing angle. In this calculation a ‘pancake’ model of the Galaxy was assumed. It was assumed that the Galaxy was a cylinder with radius 16.2 kpc and height of 1 kpc, with the location of the Sun being 8 kpc from the Galactic centre. This model was implemented by Dr. Paul Hancock. We can calculate the distance to the edge of the Galaxy for some position measured in Galactic coordinates and converted to an angle in radians:

$$D = \min \left\{ \begin{array}{l} r_S \cos \theta + \sqrt{r_G^2 - r_S^2 \sin^2 \theta} \\ \frac{h_G}{2} \frac{1}{\sin \phi} \end{array} \right. \quad (2.6)$$

where r_S , r_G , h_G are the radius to the Sun, radius of the Galaxy and the height of the Galaxy respectively. θ and ϕ are the longitude and latitude galactic coordinates of a sources’ position. We then assume that scattering screen is halfway between us and the edge of the galaxy and so halve the calculated distance (D). The scattering screen distance is then used to calculate the Fresnel scale (Equa-

tion 1.18).

$$r_{\text{F}} = \sqrt{\lambda D / 2\pi}$$

Lastly the program can also use r_{diff} to calculate the scattering disk size (θ_{scatt}) for a given line of sight (Equation 1.24).

$$\theta_{\text{scatt}} = \frac{\lambda}{2\pi r_{\text{diff}}},$$

where λ is the wavelength of the observation. Originally the SM2017 basis just did point source modulation calculations (Equation 1.21) but it has now been modified so that it can calculate extended sources if need be. As described in Section 1.4; any source smaller than θ_{scatt} will be scattered and any source larger will be down modulated based upon the ratio of θ_{scatt} to θ_{S} . We can calculate the extended source modulation index (m_e ; Equation 2.7) as well as the extended sources timescale (Equation 2.8).

$$m_e = \xi^{-1/3} \left(\frac{\theta_{\text{scatt}}}{\theta_{\text{S}}} \right)^{7/6} \quad (2.7)$$

$$t_{\text{scint}} = t_{\text{ref}} \frac{\theta_{\text{S}}}{\theta_{\text{scatt}}}, \quad (2.8)$$

where m_e is the extended source modulation index, θ_{S} is the source size, θ_{scatt} is the scattering disc size and t_{scint} is the scintillation timescale. Similarly to SM2017, we can output the data to a table for further analysis. This version of the program was then used in Hancock et al. (2019) to determine correlation between H α and observed modulation.

2.3.1 Results from RISS19

Using the RISS19 model Hancock et al. (2019) are able to determine if there is any correlation expected between the H α measured and the scattering found by a low frequency survey using the Murchison Widefield Array. The authors showed that RISS19 could be used as an all-sky model to predict variability on a variety of timescales, survey locations, and observing frequencies. Hancock et al. (2019) do this by averaging over the longitude across their data and investigating if there are any trends. One might expect that there would be more variability

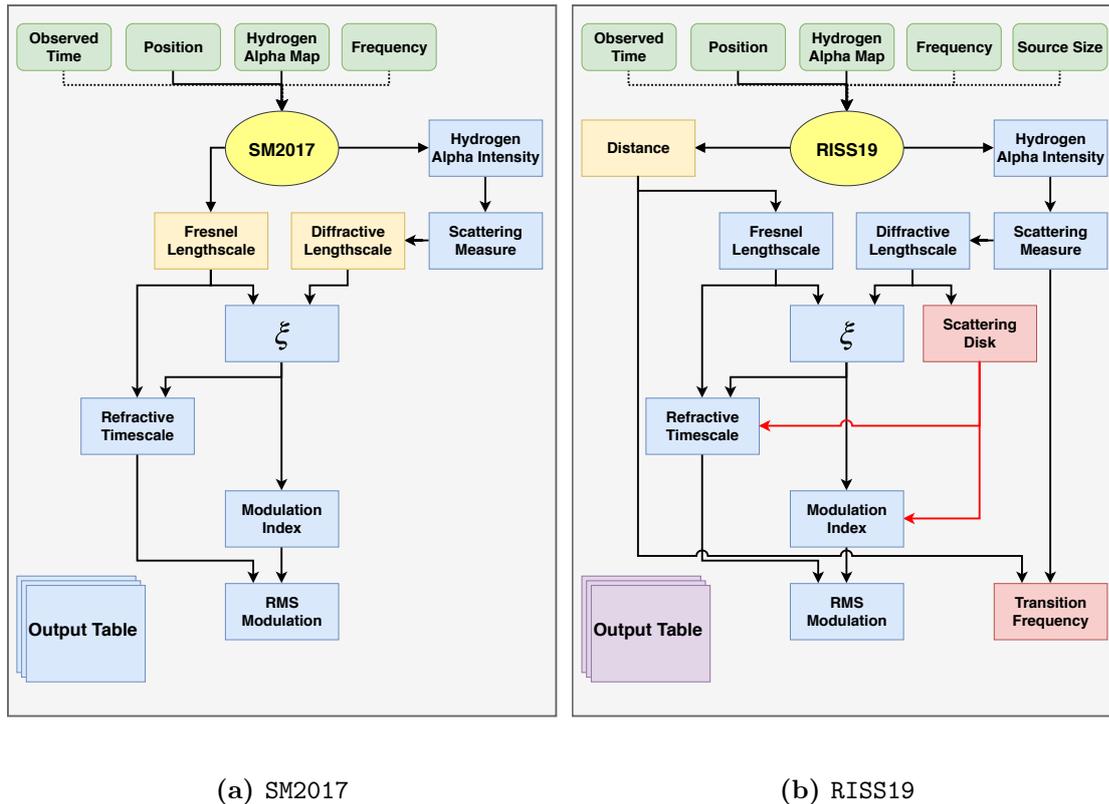


Figure 2.1: (a) SM2017 and (b) RISS19 program flowcharts. Green boxes input data and the dotted connectors represent parameters which have a default value but can be changed manually if desired. Yellow, blue and red boxes show calculated outputs where blue and red can be outputted into an output table. red boxes indicate new parameter calculations included within RISS19. The scattering disk parameter is only considered in other calculations when a source size is given to the program initially.

towards the Galactic plane where the $H\alpha$ intensity is higher. As we look through the galactic plane, we encounter more and more scattering material along the line of sight. Each scattering/scintillation event modulates the observed flux. An observer would see the sum over a large number of these events which interact with one another, and this results in a small total modulation to the observed flux (Figure 2.2). Further the paper makes a qualitative comparison between a low frequency survey here on referred to as the ‘High Galactic Latitude’ (HGL) Survey (Hancock et al., 2019) and the predicted scattering for each source position within the survey. The HGL survey observes 4 circular fields ranging in latitude from -16 to -90 degrees (Galactic coordinates) with a radius of 15 degrees (Table 2.3).

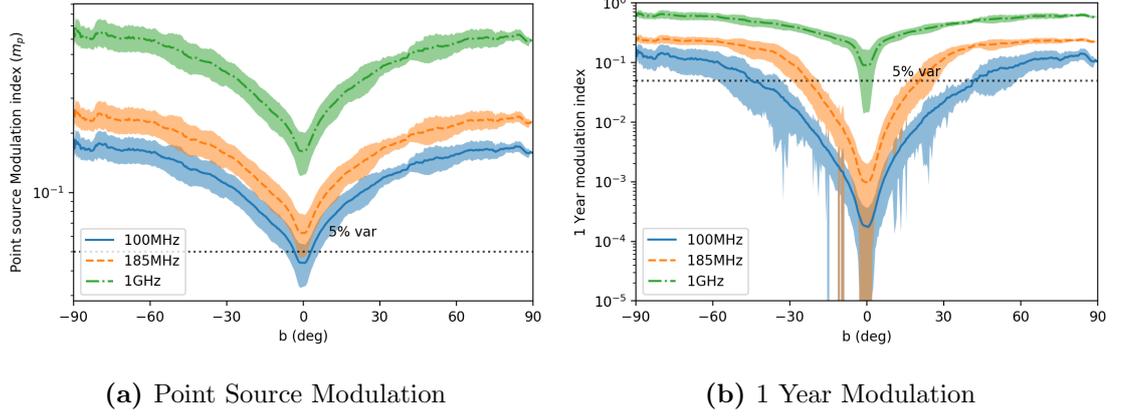


Figure 2.2: Figure showing the average total modulation for a point source (a) and the modulation expected on a one year period for a point source (b). The three different frequencies displayed are: 100 MHz (blue), 185 MHz (orange) and 1 GHz (green). The solid line represents the average value and the shaded cells are the standard deviation. Figures from [Hancock et al. \(2019\)](#).

Field	RA (deg)	Dec (deg)	ℓ (deg)	b (deg)
5	24	-24	198	-80
6	50	-20	209	-55
7	65	-13	208	-39
8	85	-1	205	16

Table 2.3: The positions of the HGL fields. Each field has a radius of 15 degrees.

Variables were selected based on two criteria: $m_d > 0.05$ and $P < 0.001$, where m_d is the de-biased modulation index which weights the modulation index by the flux uncertainties and P is the probability of seeing the observed light-curve of a source given that there is no variability. They found that 59 sources were variable, with 45 sources matching an object in SIMBAD ([Wenger et al., 2000](#)). [Hancock et al. \(2019\)](#) then plot the predicted variability against the measured variability of each source and they find that the RISS19 values over-predict the actual results. However, the measured data follows the expected trend for sources with higher $H\alpha$ to have decreased modulation, as seen in [Figure 2.4](#). At best the RISS19 model provides an upper limit to the variability. It is the aim of this

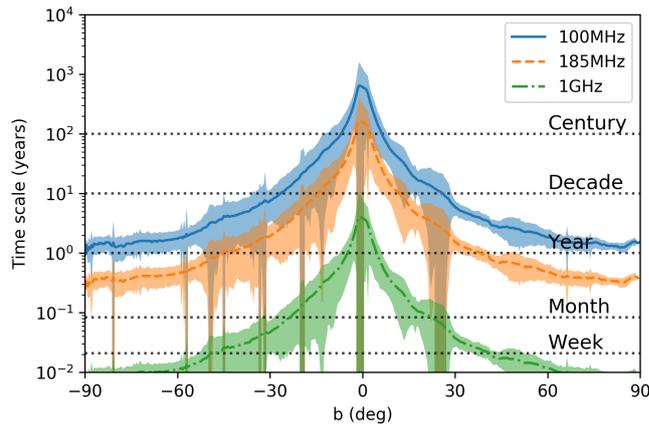


Figure 2.3: Figure showing the timescale of modulation of refractive scintillation for each latitude at each frequency. The three different frequencies displayed are: 100 MHz (blue), 185 MHz (orange) and 1 GHz (green). The solid line represents the average value and the shaded cells are the standard deviation. Figure from [Hancock et al. \(2019\)](#).

project to improve the predicted results of RISS19. [Hancock et al. \(2019\)](#) notes some limitations (L1-4):

- L1 The fractional variation of electron density and the outer scale of turbulence are assumed to be the same independent of pointing direction. These variables are critical for calculating the scattering measure ([Equation 1.15](#)). Neither of these variables can be measured directly and thus commonly accepted values are used.
- L2 RISS19 models the Galaxy as a cylinder of uniform electrons, which is used to calculate distance to the scattering screen. However, we know that the Galaxy is not uniformly distributed and there are under and over densities of electrons, with spatial variations.
- L3 The [Finkbeiner \(2003\)](#) $H\alpha$ maps used are useful due to their all sky coverage however, this also comes at the cost of large uncertainties in poorly mapped areas.
- L4 This model assumes that ISS is the only contribution however, at different frequencies ionospheric and interplanetary scintillation ([Morgan et al., 2017](#)) can contribute towards the total variability. In theory these scintillations can be separated due to different characteristic timescales.

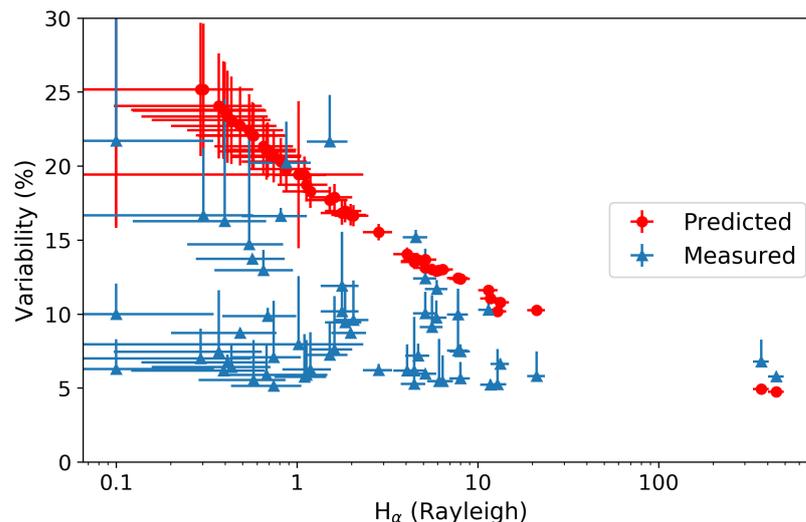


Figure 2.4: Variability versus H_α . This plot shows the HGL measured data (blue triangles) and the trend of reduction in variability at higher H_α as predicted by *RISS19* (red circles). Figure from [Hancock et al. \(2019\)](#).

We note the following solutions (S1-4) as proposed by [Hancock et al. \(2019\)](#):

- S1 The commonly accepted values of fractional variation of electron density and the outer scale of turbulence should change as a function of position. So rather than using just a constant value instead a variable or equation that changes as a function of sky position would provide much better approximations.
- S2 Use pulsar and FRB measurement models such as *NE2001* or *YMW16* to create a weighted screen distance based on the electron density, or to remove the dependence on distance altogether through the ‘Bhat Relation’.
- S3 Obtaining all sky data from a single instrument will reduce those uncertainties within the model that are a result of poor input data.
- S4 The contributions of the different regimes of scintillation and scattering should be accounted for. However, MWA images taken over two minutes would only see an increase of detected variability of approximately 1% due to interplanetary and ionospheric scintillation.

In this project we aim to address L2 and L3 by applying S2 and S3 ([Section 2.4](#)). L1 will not be addressed as there is a lack of data across the whole sky to mean-

ingly contribute to this project. L4 is in relation to two different mediums (ionosphere/IPM) which occur on different timescales and would require a whole new set of models, which cannot be addressed reasonably in the timeline of this project.

2.4 Improvements to RISS19

2.4.1 Hydrogen Alpha Maps

The first limitation (L3/S3) that can be solved is the issue of the [Finkbeiner \(2003\)](#) maps having large uncertainties by using instruments with smoother coverage across the whole sky. In this research $H\alpha$ is being used as a proxy for scattering measure ($I_{H\alpha} \rightarrow EM \rightarrow SM \rightarrow r_{\text{diff}}$) which allows us to determine how much scintillation we should see for a given line of sight (see [Subsubsection 1.3.2.1](#)). Therefore it is important that this data is as accurate as possible across the whole sky.

The initial $H\alpha$ map used was a composite map made by [Finkbeiner \(2003\)](#) using three surveys. The three surveys were: Wisconsin Hydrogen Alpha Mapper (WHAM) [Haffner et al. \(2003, 2010\)](#), Virginia Tech Spectral-line Survey (VTSS) [Dennison et al. \(1998\)](#), and The Southern H-Alpha Sky Survey Atlas (SHASSA) [Gaustad et al. \(2001\)](#). The Finkbeiner map created from this was very useful because it had large sky coverage with a FWHM resolution of 6'. However, there were some issues with this map. One of the major issues is that there was very little coverage in the southern hemisphere, and so the data in southern hemisphere regions had large uncertainties, which were several times larger than the actual $H\alpha$ intensity values. Also where each survey is stitched together there are large uncertainties due to change in resolution and surface brightness sensitivity of each telescope. The uncertainties at the poles are also quite large ($\sim 100\%$) and are stretched due to projection mapping; the top (90°) and bottom (-90°) rows of the 2D projection are actually just one pixel on a 3D sphere that has been stretched.

All Sky WHAM Map

To date no all sky $H\alpha$ maps from a single instrument have been published. However, the WHAM team that originally mapped the northern hemisphere moved their telescope to the southern hemisphere in 2010. The WHAM sky survey has recorded data and an early set of data was released (Haffner et al., 2010). Although, the full sky survey paper is still not published, the full sky survey data is nonetheless available. This data has a resolution of one degree across the whole sky, with significantly less errors as seen in Figure 2.5b.

The available data included the fits table and image for the $H\alpha$ observations, although no processed uncertainty map had been produced. We decided to re-create the $H\alpha$ image and produce an uncertainty map from the fits table provided. This was done by using `scipy.interpolate.griddata` to interpolate points, using the ‘linear’ interpolation method in combination with ‘nearest neighbour’ to mask `nan` values. This map will be referred to as the All Sky WHAM map (AS-WHAM) from here on. The comparison between the Finkbeiner and the AS-WHAM maps can be seen in Figure 2.5a we can see the distribution of the $H\alpha$ with its variance, and in Figure 2.5b we can see the significant difference in the fractional uncertainties between the two maps.

Further, Figure 2.6b shows the reduction in uncertainties when the results from Hancock et al. (2019) are re-calculated and re-plotted. A comparison of the percentage error maps can be seen in Figure 2.7, it is clear that the AS-WHAM map has significantly lower uncertainty across the whole map, with the majority of the pixels having an error below 20%. This map eliminated the composition issues caused by stitching together three separate maps.

However, this does come at the cost of resolution, and while the Finkbeiner map has a resolution of 0.1 deg, the AS-WHAM map has a resolution of approximately 1 deg. The difference in resolution can be seen in Figure 2.8, where some of the fine structure is lost. I decided that the trade-off between largely decreased

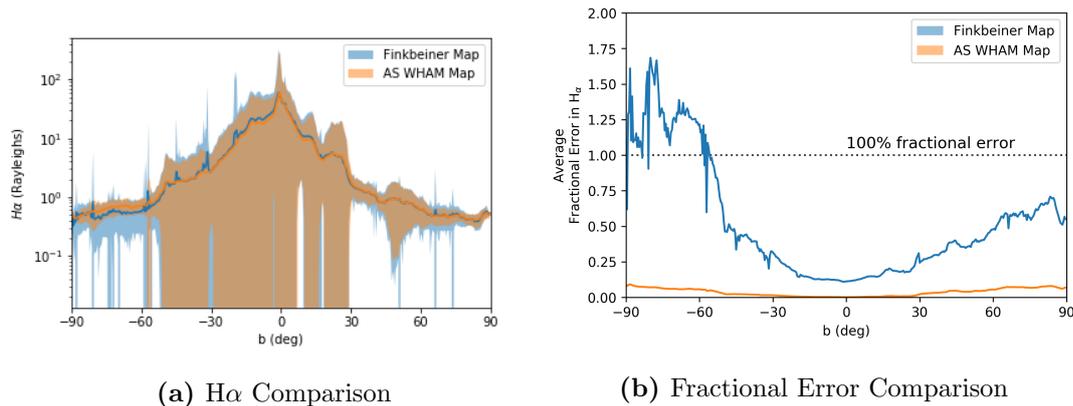


Figure 2.5: Left: comparison of the averaged (over the longitude) $H\alpha$ distributions with the variance shown as the shaded parts of each line. Right: The distribution of average (over the longitude) fractional errors in the $H\alpha$ measurements. In both plots the orange is the All-sky WHAM map, and blue Finkbeiner map. Adapted from [Hancock et al. \(2019\)](#).

uncertainties is worth the small loss in resolution.

2.4.2 Distance Models

The second limitation (L2/S2) we can address is the simplistic model of the Galaxy, by accounting for under/over densities of electrons or by removing dependence on the distance all together. In [RISS19](#) the distance to the scattering screen (D) is an important input parameter, as several parts of the program are reliant on this measurement, the two important outputs being the modulation index (m_p) and the scintillation timescale (t_{scint}).

$$m_p = \left(\frac{r_F}{r_{\text{diff}}} \right)^{-1/3} \propto (\sqrt{D})^{-1/3} \propto D^{-1/6} \quad (2.9)$$

$$t_{\text{scint}} \propto D \quad (2.10)$$

The distance to the scattering screen was modelled through several different iterations. The first being a set distance of 1 kiloparsec; this is a very oversimplified model of the distance which does not take into account our position within our Galaxy or the distribution of scattering material along the line of sight. The

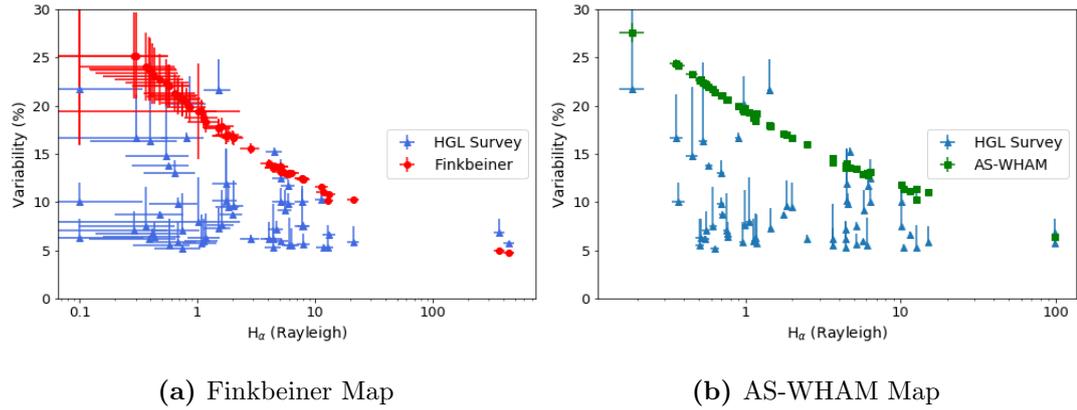


Figure 2.6: (a) Shows the results from [Hancock et al. \(2019\)](#). (b) Plot showing the results from [Hancock et al. \(2019\)](#) recalculated using the new $H\alpha$ map.

second iteration modelled our Galaxy as a pancake (i.e. a thin cylinder) and calculated the distance to the edge of the Galaxy and then placed the scattering screen at the halfway point of this distance. This was also too simple and does not take into account the densities of certain regions contributing to the scattering measure such as the Galactic centre. The newest iteration used the **YMW16** model ([Yao et al., 2017](#)) which models the electron content and distribution of different parts of the Galaxy; Galactic centre, spiral arms, thin disk, thick disk, we called this the electron density model.

Electron Density Model

In the electron density model we want to weight the position of the scattering screen based on the density fluctuations of electrons along a line of sight. However, as mentioned in [1.3.2.1](#) this can be hard to measure, instead we say that these fluctuation are proportional to the electron density square. The **YMW16** model can return DMs and electron density at a given position and distance, this allows us to construct a map of the n_e^2 along a given line of sight.

[Figure 2.9](#) shows the distribution of electron density down given lines of sight. It also shows the three different models used for the screen distance. The first model is the 1 kilo-parsec screen in blue, this is an often assumed distance to the scattering screen and is the most basic model. The second model is the pancake

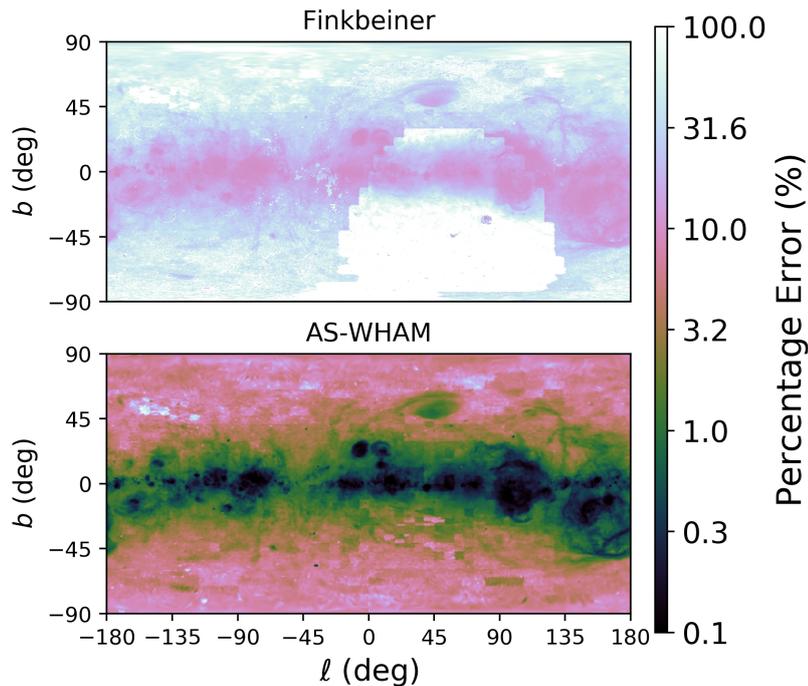


Figure 2.7: The percentage error of the Finkbeiner (top) and AS-WHAM (bottom) Map. The flux scale ranges from 0.1% (black) to 100% error (white).

model in green used by RISS19. This models the Milky Way as a pancake and places the screen halfway between us and the edge of the Galaxy. The third and final model is the weighted fit (red). This model weights the position of the screen based on the fluctuation in the density of the electrons (n_e^2). The weight model was chosen as the final model as it incorporates some aspects of the structure of the Milky Way. While this weighted model is a much better approximation of the scattering screen distance, we could improve the model further by removing the dependence on distance.

Tau Model

Measurements of pulse temporal broadening (from here on out called ‘tau’ - τ_d) from pulsar observations are found to be related to DM, via the so called ‘Bhat Relation’ (Bhat et al., 2004). This allows for the removal of the explicit dependence on the scattering screen distance by relating DM to τ_d to obtain the amount of scattering directly. We will call this the ‘Tau Model’. As all of our

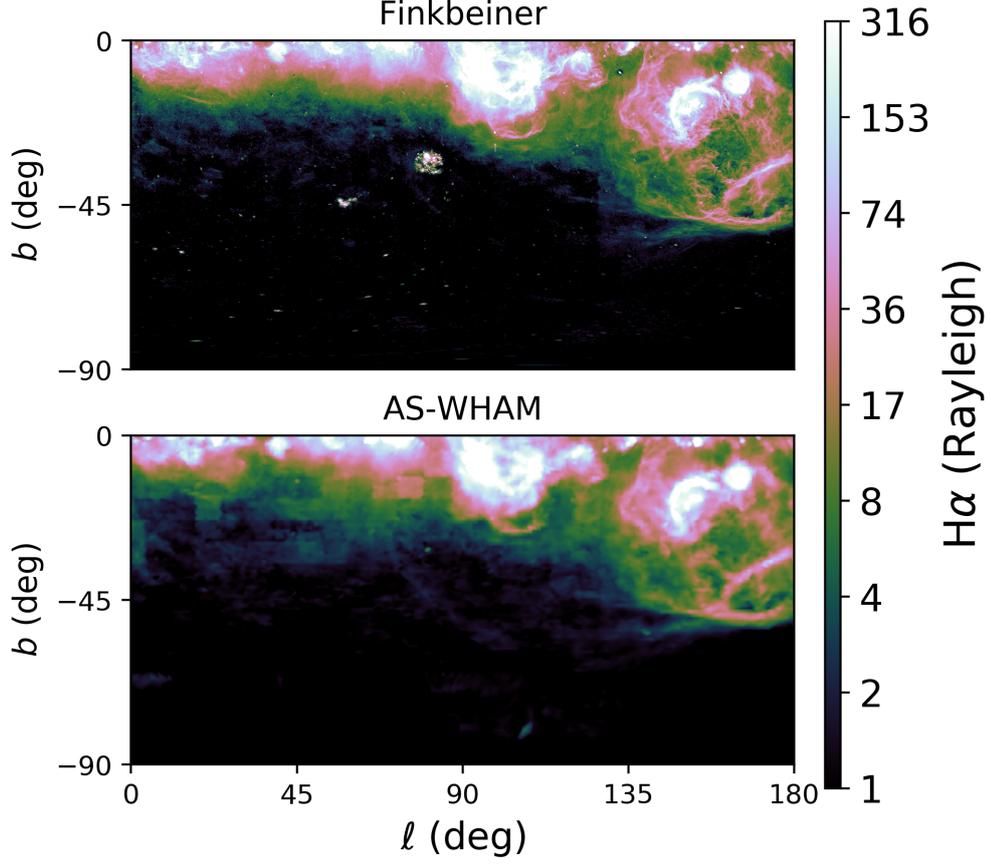


Figure 2.8: The $H\alpha$ maps from the Finkbeiner (top) and AS-WHAM (bottom) data. The flux scale ranges from 1 Rayleigh (black) to 316 Rayleigh (white). Due to the authors selection of velocity ranges, the Magellanic Cloud is lost in the AS-WHAM data.

sources are extra-Galactic, we can assume that all the electrons between us and the edge of the Galaxy contribute to the DM and the scintillation. The YMW16 model use the Bhat Relation so that it can be used to calculate DM and τ_d value. The relation for the pulse delay time (τ_d) in seconds to DM used in the YMW16 model is given by:

$$\tau_d = 4.1 \times 10^{-11} \text{DM}^{2.2} (1.0 + 0.00194 \text{DM}^2). \quad (2.11)$$

This is normalised to 1 GHz with an assumed scaling factor of $\tau_d \propto \nu^{-4}$. We now want to relate τ_d to a quantity that is useable in our program such as the diffractive and Fresnel scale, which we do by using some substitutions to find:

$$\tau_d = \frac{D_{\text{eff}} \theta_{\text{scatt}}^2}{2c} \quad (2.12)$$

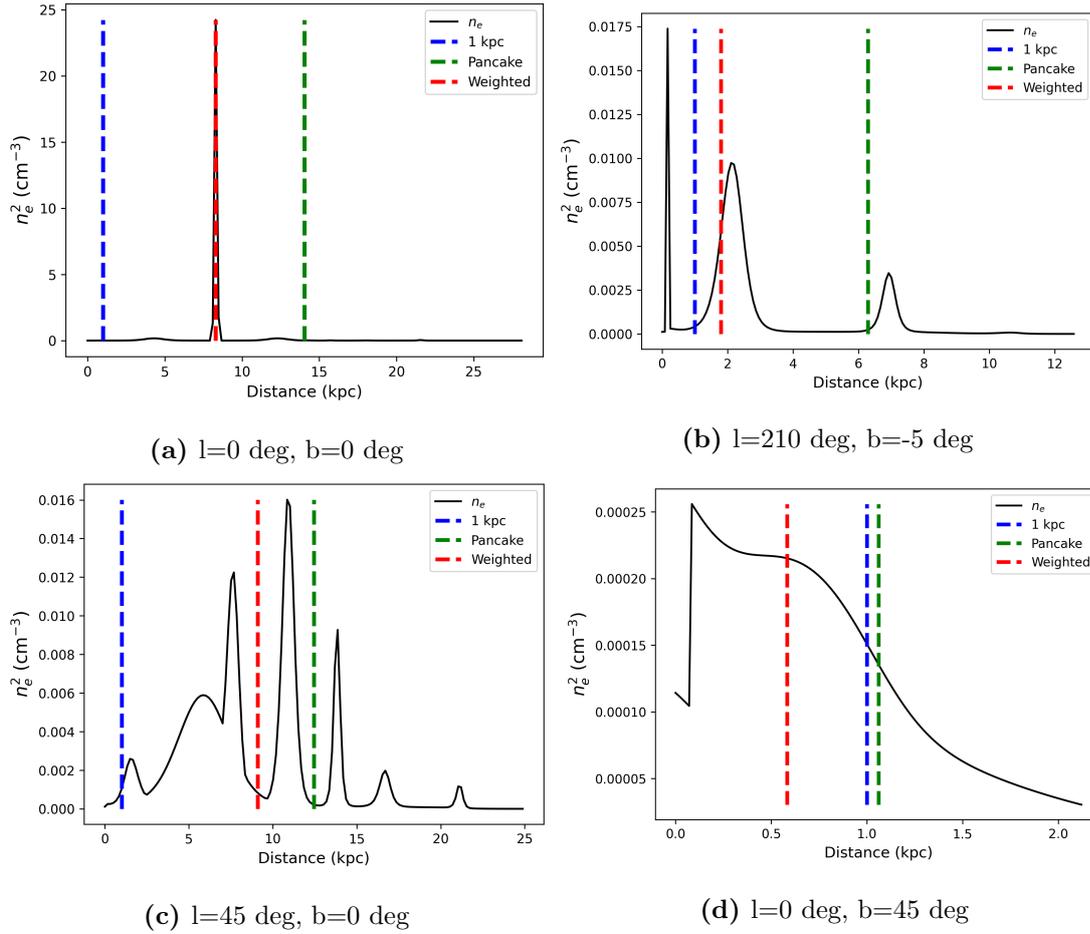


Figure 2.9: Figures showing the distribution of n_e (black) down a line of sight, with three different models for weighting the screen distance. The different distance models are: 1 kpc (blue), ‘pancake’ (green) and the weighted fit (red). The 3 different models are shown for 4 lines of sight.

$$\theta_{\text{scatt}} = 1/kr_{\text{diff}} \quad (2.13)$$

$$r_{\text{F}} = \sqrt{\frac{\lambda D}{2\pi}} = \sqrt{\frac{D}{k}}$$

$$\tau_{\text{d}} = \frac{1}{4\pi\nu} \left(\frac{r_{\text{F}}}{r_{\text{diff}}} \right)^2 \quad (2.14)$$

We can then use this to recalculate the equations for the other quantities that the RISS19 model calculated. The quantity r_{diff} is still inferred using the Hydrogen alpha map.

$$m_{\text{p}} = (r_{\text{diff}}/r_{\text{F}})^{1/3} \quad (2.15)$$

$$m_{\text{p}} = (4\pi\nu\tau_{\text{d}})^{-1/6} \quad (2.16)$$

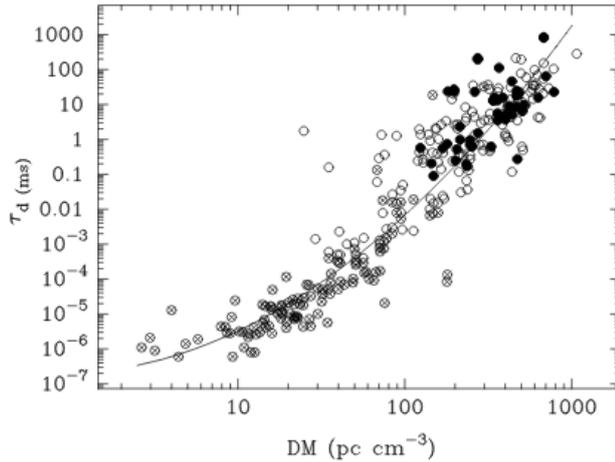


Figure 2.10: The Bhat relation (Bhat et al., 2004). Here we can see the pulse delay in milliseconds versus the dispersion measure. Filled circles are measurements made by Bhat et al. (2004), crossed circles are derived from decorrelation bandwidths and open circles are other published pulsar measurements. The solid line represents the best-fit model for this relation.

The point source modulation index is no longer reliant on $H\alpha$ Equation 2.16, instead it is solely determined by the YMW16 model. However, the extended sources are still dependent on $H\alpha$ as can be seen in Equation 1.24 as r_{diff} is calculated using the $H\alpha$ intensity. Similarly we can write r_{F} in terms of τ_{d} and r_{diff} :

$$r_{\text{F}} = r_{\text{diff}} \sqrt{4\pi\nu\tau_{\text{d}}}, \quad (2.17)$$

this allows us to then recalculate all of the the RISS19 parameters in terms of τ_{d} instead of distance. It should be noted that the majority of known pulsars are within the Galactic plane and so the map has much better predictive power in these regions when compared to positions off the Galactic plane (Figure 2.11), where significant interpolation is required. As FRBs are extra-Galactic they are not bound to being discovered within the plane. With the increasing number of FRB searches, detections and localisations using ASKAP and CHIME (Bannister et al., 2019; CHIME/FRB Collaboration et al., 2019), models such as YMW16 will only become more accurate in their predictions. The τ_{d} map can be seen in Figure 2.12

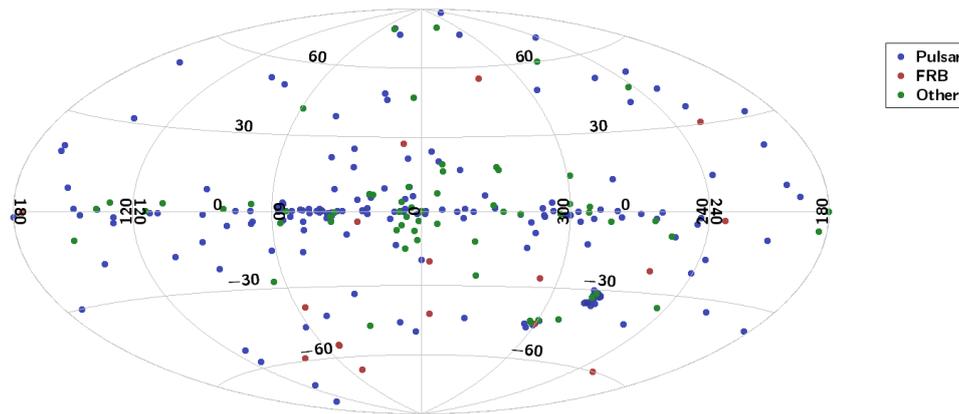


Figure 2.11: Plot of the pulsar (blue), FRB (red) and other galactic components (green) used to generate the YMW16 model, plotted in Galactic coordinates as an aitoff projection. Data obtained from Yao et al. (2017).

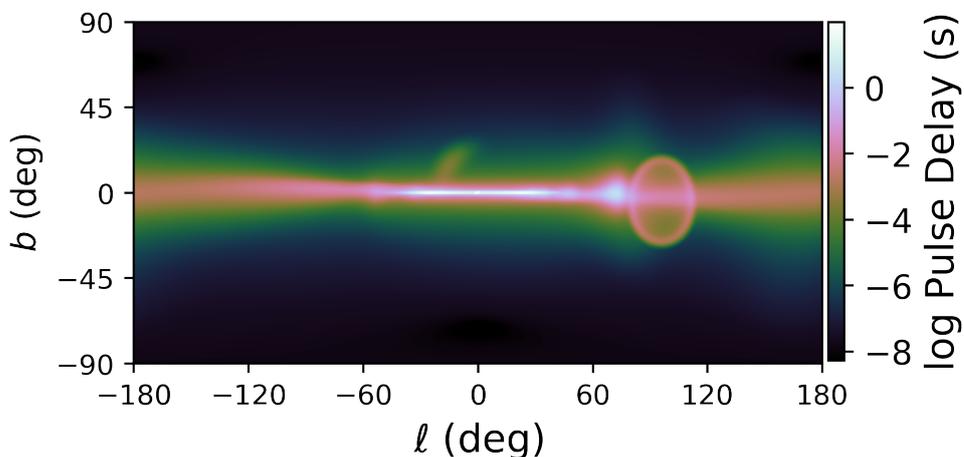


Figure 2.12: The map of the pulse delay in seconds, plotted as a function of Galactic latitude (l) and longitude (b).

2.5 Results

We have now updated the RISS19 model, and addressed some of the limitations noted in Hancock et al. (2019). However, even with updates, the model still has limitations. Firstly, the model is still heavily dependent on screen distance, and even though we have iterated to a better distance model it is still not perfect. Future updates could be easily implemented as further observations are made and incorporated into updated models. We improved the $H\alpha$ map by sourcing more complete data from a single instrument. However, $H\alpha$ itself is not a perfect solution to predicting electron density throughout the Galaxy, as it is not

a direct measure of SM. Solutions to this would come in the form of combining several different tracers for the ISM, such as carbon monoxide and other atomic lines. We have started this process by using $H\alpha$ and DMs (TAU19), but further development lies outside the scope of this project. Furthermore several limitations stated in [Subsection 2.3.1](#) have not been addressed as they were either not able to be incorporated into the project's timeframe, or there was not sufficient data to complete them. However, these limitations should still be remembered when using the model.

In this section we discuss results for one of the project's goals: to simulate variability for given positions across the sky. To do this we will first use the HGL survey. In this section we will use the new AS-WHAM $H\alpha$ map by default.

2.5.1 Comparison of Distance Models

[Figure 2.13](#) shows the comparison of variables found by [Hancock et al. \(2019\)](#) to each new distance model at 185 MHz. When comparing this figure to [Figure 2.2](#), we can see that the RISS19 and the new weighted distance models (WD-RISS19) models are similar to one another, with the weighted distance model having slightly higher modulation predictions due to its screen placement. The screen placement for the weighted distance model, on average, is closer and therefore increased modulation is seen ($m_p \propto D^{-1/6}$). On the other hand the TAU19 model predicts less modulation for a point source, but for the 1 year modulation it predicts a significantly decreased modulation on average. This is likely due to lines of sight near the Galactic plane having a large predicted electron density, therefore increasing the scattering time. Additionally as the TAU19 model is based on pulsar measurements, the majority of which are in the galactic plane, the model is dominated by lines of sight with large pulse delays. From this comparison and previous discussion, the new weighted distance model will be our default model over the original RISS19 model, from here-on the RISS19 model shall use the weighted distance model. We will further compare RISS19 and TAU19 against actual survey results.

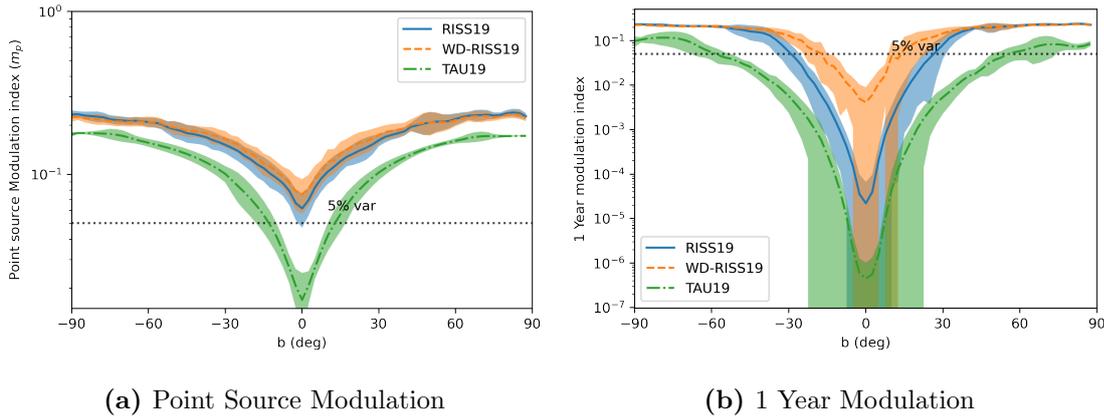


Figure 2.13: Figures comparing each distance model: **(a)** point source modulation **(b)** scaled modulation on 1 year timescales. The three different models displayed are: RISS19 (blue), WD-RISS19 (orange) and TAU19 (green). The solid line represents the average value and the shaded cells are the standard deviation. Figures adapted from [Hancock et al. \(2019\)](#).

2.5.2 Comparison to HGL

The High Galactic Latitude (HGL) survey ([Hancock et al., 2019](#)), is a 185 MHz survey created using the MWA. It contains 4 fields each decreasing in latitude. This survey found 58 sources that were classified as variable. We can use the positions of these sources in combination with our model to generate modulation statistics. In [Figure 2.14](#) we can see the comparison between the observed data (blue), the RISS19 model (orange), and the TAU19 model (green). Generally both models over predict the data, however, this is expected as the model assumes the sources are point like, and that the timescale of modulation is saturated. This means that the model will be predicting the maximum amount of variability, and so we expect this model to act as an upper limit. The TAU19 model appears to better fit the data, however, this is the maximum modulation the TAU19 model will predict. If we instead use the actual time frame of the observations (448 days) then the TAU19 significantly under predicts the data ([Figure 2.15](#)), due to the timescale dependence seen in [Figure 2.13b](#). Since the RISS19 model scales less steeply with time, a much better agreement than the TAU19, while still over predicting the variability when compared to the survey results. Neither program

is incorrect, however, the RISS19 model provides a good upper limit that scales as expected from the theoretical background. Based on these results, we will use RISS19 as the default model moving forward.

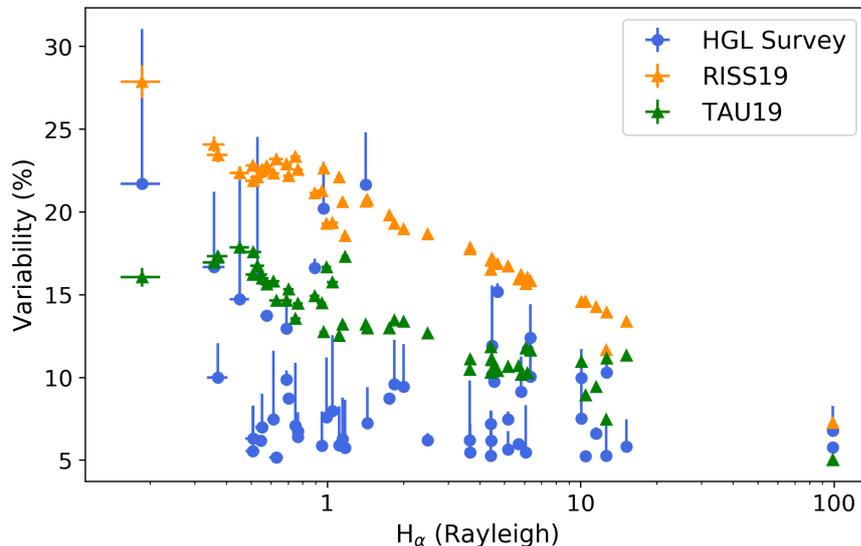


Figure 2.14: Comparison of the expected modulation of the different scintillation distance models against observational variability data from the HGL Survey (blue circles). RISS19 model is described by orange triangles, and the predictions of TAU19 model by green triangles.. Figure adapted from [Hancock et al. \(2019\)](#).

2.5.3 Multi-Frequency Data

Another example of use for the RISS19 is to predict the amount of variability scintillation could account for across multiple frequency bands. Such a use case arose during my colleagues (Ms. Kathryn Ross) PhD project. They are detecting and classifying spectral variability at low frequencies between two data sets captured one year apart by the MWA. Using the provided data by Ms. Ross in combination with RISS19 we put an upper limit on how much scintillation was contributing to these variations.

This can be seen in [Figure 2.16](#) where we have the first use: spectral energy distribution (SED), for one of these sources. The two sets of data taken one year

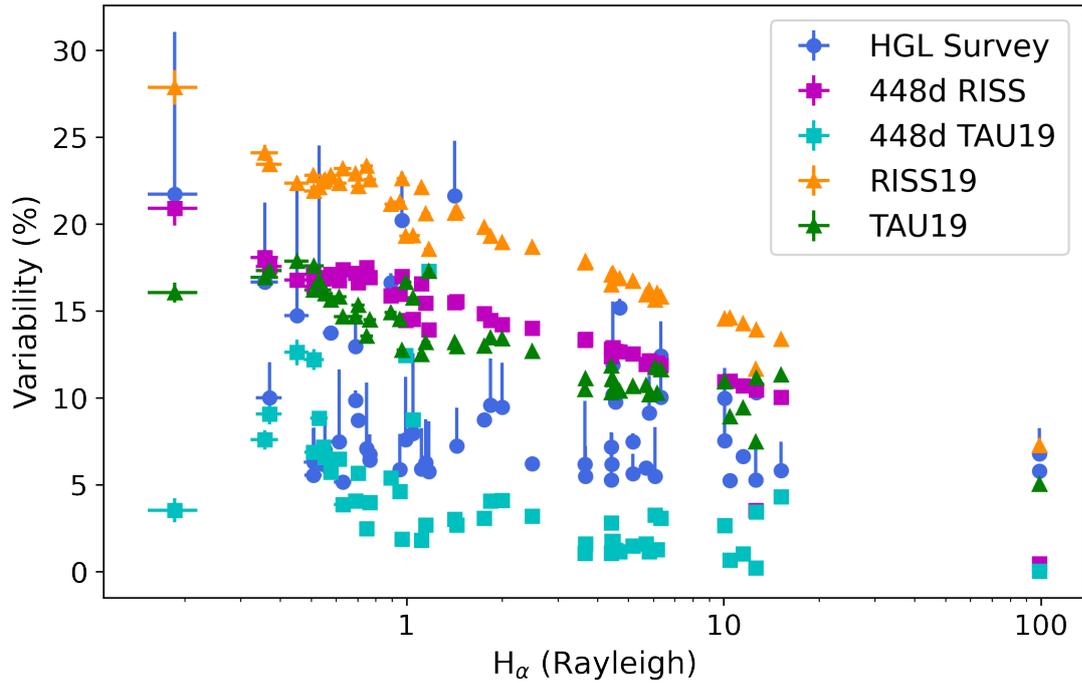


Figure 2.15: (a) Comparison of the expected modulation of the different scintillation distance models against observational variability data from the HGL Survey (blue circles). The predictions of the RISS19 model at its full modulation are shown by orange triangles, and at 448 days by magenta squares. The predictions of TAU19 model are shown by green triangles, and at 448 days by cyan squares. (b) The change in variability between the 448 day predicted modulation (square markers) and the maximum modulation (triangle markers), RISS19 is shown in orange and TAU19 is shown in green. Figure adapted from [Hancock et al. \(2019\)](#).

apart (black and green points) are contrasted with the prediction for one year modulation for the RISS19 (blue). From this figure we can see that the lower frequency points (~ 100 - 150 MHz) variability could be explained by scintillation, however, the higher frequency points could not be entirely explained by scintillation (at least not within one sigma errors). These results can be used to help an observer assess to what extent their sources' variability could be attributed to interstellar scintillation.

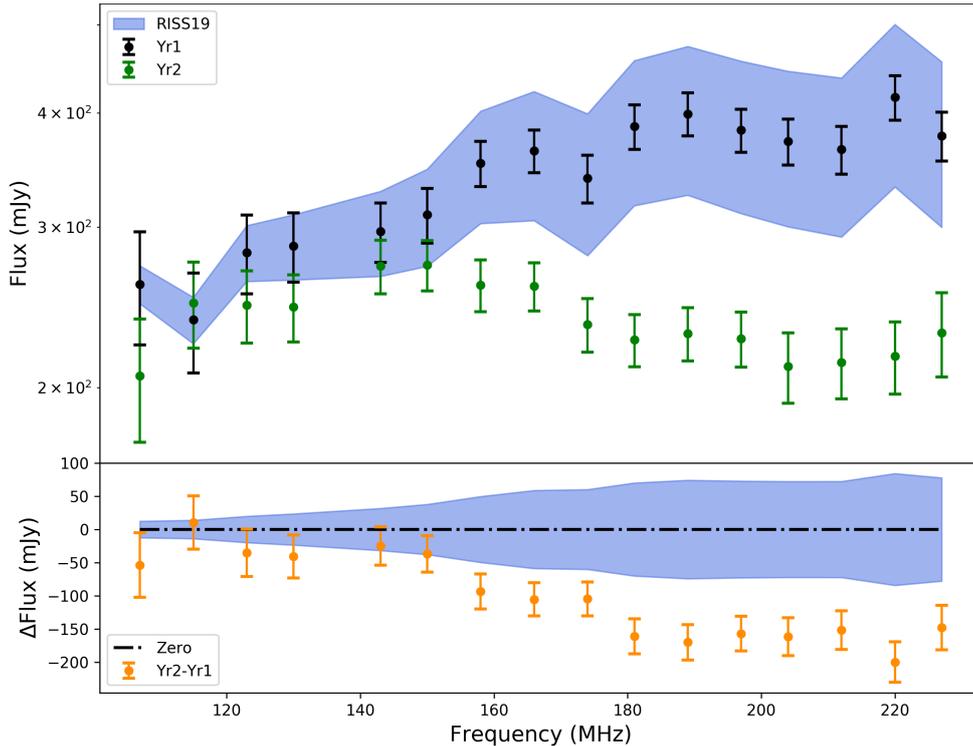


Figure 2.16: Top: Spectral energy distribution for one of Ms. Ross’ sources. This diagram shows the log flux density (mJy) versus the frequency (MHz). In blue is the first year data, in green is the second year data. Further in blue shaded areas we have the predicted modulation of flux from the RISS19. Bottom: Residuals for the SED, the data shown is: zero line (dot-dashed black line), year 2 - year 1 (orange circles) and the modulation predicted by RISS19 (blue shaded area).

2.6 Summary

In this chapter we have described our updates to the RISS19 model to address for some of the limitations identified by Hancock et al. (2019). These imitations specifically include poor H α map quality and simplistic distance models used. We corrected these by sourcing complete H α maps from a single source (Haffner et al., 2010) with significantly reduced uncertainties. Furthermore we iterated to a new distance model that took into account the structure of our Galaxy using the Yao et al. (2017) model for the Galaxy. While these limitations were addressed we noted that further improvements could still be made to the model by incorporating: different tracers for the ISM to make for better electron density predictions, an iteration on the distance model once more data and better models

become available, as well as limitations stated by [Hancock et al. \(2019\)](#) that could not be addressed due to insufficient data at the current time or being outside the timeframe of the project.

From these updates we have then tested the two distances models together and their predicted modulation for each model. From this we chose to further test the weighted distance model of **RISS19** and the **TAU19** model. We then used these two models to predict line of sight variability for the HGL observations. We found that **RISS19** presents an upper limit on the expected variability, whereas **TAU19** seems to fit the data better. However, upon application of the surveys actual observation time (448 days) we found that the **TAU19** model change from over-predicting by $\sim 10\%$ to under-predicting by $\sim 10\%$, whereas the **RISS19** model had better agreement it still acted as an upper limit to the variability. From this we decided that the **RISS19** model was more stable and would be best used as the main model. Finally we tested the **RISS19** model for multi-frequency data, to predict a sources variations across frequency. We used the model to give a limit to the amount of modulation in the flux one could expect from each sources' line of sight. This application of the model allows one to quantify the scintillation in comparison to the variability observed.

We have found that the model works and can produce upper-limit results for single or multiple lines of sight, as well as across different frequency bands. Thus meeting one of the key goals of the project. Now we can use this model in combination with information about sources and their distribution to simulate a region of sky. This is our final use case; to simulate variability statistics for regions of the sky before an actual survey is taken. In [Chapter 3](#) we will describe how this simulation works, the results, and limitations found.

Chapter 3

Simulation and Surveys

3.1 Introduction

In the previous section we described the **RISS19** model, which predicts the modulation index and other variability parameters for given lines of sight. We also have data from Dr Paul Hancock and public sources on observed variability studies. We want to compare these two data sets to see if our scintillation model holds up or not. Doing this using various surveys can also allow us to work out why or where the model fails, which can provide insight into improvements that can be made. In order to simulate a survey we require: the sources that are being observed, the source properties, and the region of the sky that the population belongs to. So to compare our predicted variability to that of observed surveys, we need to create a program that can generate realistic sources to be used in combination with the **RISS19** model to generate variability trends for the population.

We first restrict the simulation to only include extra-galactic radio sources and exclude any Galactic sources. There are several properties of these sources that need to be generated. Properties such as size, structure, flux, and position of the source, will each affect the degree of predicted variability. An additional difference between **RISS19** and the simulation is that while **RISS19** predicts the ‘true’ variability of the source, the simulation will predict the ‘observed’ variability. This means that properties of the telescope observing need to be accounted for,

including the telescope’s resolution, sensitivity, and operating frequency. However, source counts for compact sources, especially at low-frequencies are largely inaccurate and so this simulation will be one with large errors.

In this section we will develop a model to simulate sources within a region of the sky, then take these generated sources and use `RISS19` to calculate variability statistics about the population. We will then test our simulation and model by replicating other surveys, to test our agreement. First we shall generate sources and their properties.

3.2 Position Generation

For our simulation we want to be able to determine the number of sources within a given region that is specified by a user, so that we can predict how many of these sources would be variable. To do this we first need to generate sources randomly across the region. However the distribution of galaxies in the Universe is not homogeneous or simple; instead the distribution is inhomogeneous, with galaxies grouping together in clusters and forming the cosmic web. Unfortunately there exists no easily implementable data to model this structure, and we don’t expect this to affect the results of the simulation so we will not include this.

We generate a uniform distribution of sources on a sphere and we use the `Region.sky_within` function from `AegeanTools` ([Hancock et al., 2018](#)) to check whether these points are within the region specified by the user. Now that the simulation generates a uniform distribution of points across the sky, we want to generate the correct number of sources at the correct flux densities. This method does have some significant drawbacks, namely that we are generating sources across the whole sky just to use sources within a certain area. This means that if we have a significantly small area (less than a degree squared) we have to generate close to a billion sources across the whole sky just to find the correct amount of sources within this small area.

3.3 Source Counts

The simulation can now distribute points evenly, we now want to generate the flux because the flux of the modulation is taken into account when determining whether the modulation seen is significant when compared to the flux of the source and the flux of the noise in the image. To do this we need to generate the correct number of sources and distribute these sources with the correct flux. This relation between number of sources (N) at a given flux (S) is called the source counts. For a steady-state Euclidean Universe the slope for this function is $\beta = 3/2$ (Equation 3.1), and so we write source counts as a power law function of $\log(N)$ vs $\log(S)$.

$$N(> S) = C_1 S^{-\beta} \quad (3.1)$$

Where N is the number of sources greater than the flux S , C_1 is a scaling constant, and β is the power law. This can also be written in the differential form (Equation 3.2) which tells us the number of sources (dN) per flux bin (dS) over a given area of the sky.

$$\frac{dN}{dS} = C_2 S^{-(\beta+1)} \text{ (Jy}^{-1} \text{ sr}^{-1}\text{)} \quad (3.2)$$

Where C_2 is a scaling constant. However theoretical approach assumes that the Universe is infinite and unchanging, however the Universe had a beginning and has evolved over time. To avoid these limitations we can instead we use multiple observational surveys which have calculated source counts. Using these pre-determined surveys would allow us to get a better estimate for the source counts at different frequencies, as seen in Figure 3.1.

Paper	Frequency	Flux Range (mJy)	Sensitivity
Franzen et al. (2019)	154 MHz	$1 \rightarrow 75 \times 10^3$	$100 \mu\text{Jy beam}^{-1}$
Hopkins et al. (2003)	1400 MHz	$0.05 \rightarrow 1 \times 10^3$	$60 \mu\text{Jy beam}^{-1}$

Table 3.1: Table shows the different surveys used, their frequencies, flux range and sensitivities for comparison.

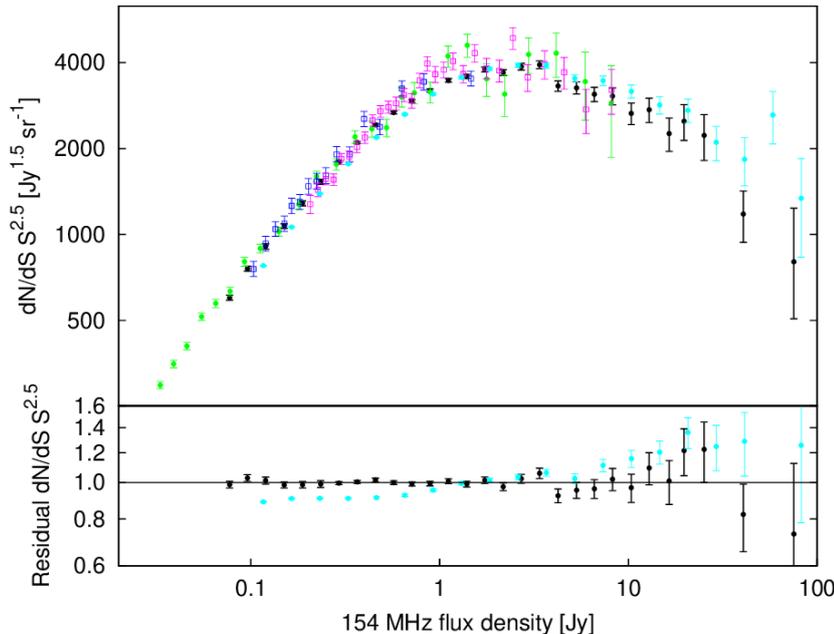


Figure 3.1: Top: The Euclidean normalised differential source counts for 150-154 MHz, where the black circles are Franzen et al. (2019), cyan circles Intema et al. (2017), pink squares Hales et al. (2007), and blue squares McGilchrist et al. (1990). Bottom shows the residuals between GLEAM (Hurley-Walker et al., 2016) and TGSS (Intema et al., 2017). Figure taken from Franzen et al. (2019).

We can use a general formula for a polynomial fit and each surveys source count polynomial coefficients to create a model for the source counts:

$$\log[(dN/dS)/(S^{-2.5})] = \sum_{i=0}^n a_i [\log(S)]^i, \quad (3.3)$$

where S is the flux in mJy, dN/dS is the number of sources per flux bin, n is the order of the polynomial and a is the polynomial coefficient at the i 'th power. Where Franzen, $a = [3.52, 0.307, -0.388, -0.0404, 0.0351, 0.006]$ and Hopkins, $a = [0.859, 0.508, 0.376, -0.049, -0.121, 0.057, -0.008]$. We can then calculate non-normalised source counts (N) from the normalised differential counts using:

$$N = \left[\frac{dN}{dS} S^{2.5} \right] \times \frac{dS}{S^{2.5}}, \quad (3.4)$$

where the singular dS is the width of the flux bin. We can plot the differential source counts and the source counts and see that the survey's differential source counts shapes are similar at their respective frequencies (Figure 3.2).

However if an observer is observing at a frequency that is not 154 MHz or 1400 MHz then we must be able to scale our surveys source counts to this frequency.

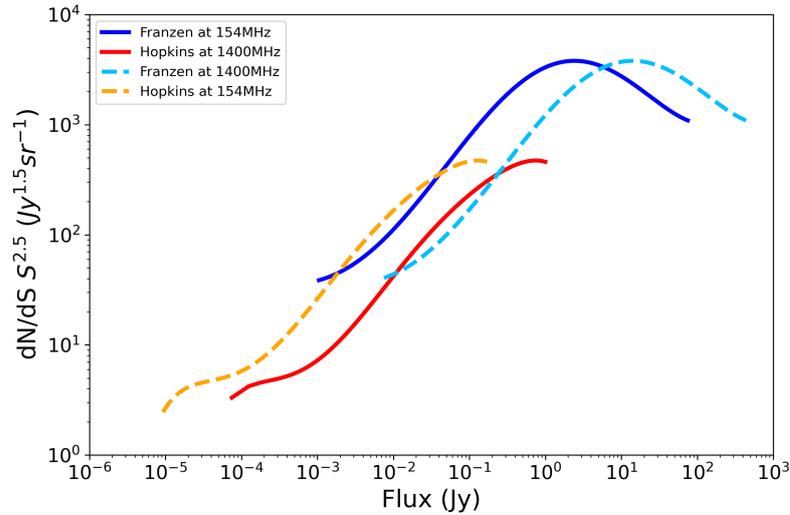


Figure 3.2: The differential source counts for each survey at their respective frequencies. In blue is [Franzen et al. \(2019\)](#) at 154 MHz and in red is [Hopkins et al. \(2003\)](#) at 1400 MHz.

We can do this by scaling our flux bins:

$$S = S_0 \left(\frac{\nu}{\nu_0} \right)^\alpha, \quad (3.5)$$

where S is the scaled flux, S_0 is the original flux, ν is the observing frequency, ν_0 is the frequency of the survey data and α is the spectral index of the source, the average spectral index for all sources found in GLEAM ([Hurley-Walker et al., 2016](#)) was $-0.83 \leq \alpha \leq -0.78$, and so an assumed value of $\alpha = -0.8$ is taken. The underlying populations of both surveys are the same and so we expect that the survey source counts will scale as the observational frequency is changed. We expect that if we observed these sources at some frequency, then re-observed them at a higher frequency we would observe a decreased flux. The opposite would be true for the sources re-observed at a lower frequency, we would see the same sources with a larger amount of flux. The model first generates source counts based on both survey's source counts and then scales the flux bin by using the given frequency and [Equation 3.5](#). This is demonstrated in [Figure 3.2](#) for both 154 MHz and 1400 MHz scaling, we can see here that the agreement between each model after scaling is good, and proves that the scaling used in [Equation 3.5](#) is consistent.

If the frequency is below 154 MHz we can assume that Franzen will model this regime better than the Hopkins counts, the same holds true for frequencies greater than 1400 MHz, Hopkins will be the default survey. However for regions in between we must apply some weighting to both surveys:

$$w_F = 1 - \frac{\Delta\nu_F}{\Delta\nu_F + \Delta\nu_H}, \quad (3.6)$$

$$w_H = 1 - \frac{\Delta\nu_H}{\Delta\nu_F + \Delta\nu_H}, \quad (3.7)$$

where w_F and w_h are weightings applied to the Franzen and Hopkins survey. $\Delta\nu_F$ and $\Delta\nu_H$ are the difference between the surveys frequency and the observing frequency. With these weightings we can then apply them to each survey to get a weighted fit.

It is important to note that when observing with a telescope there are several limitations that mean it will never find 100% of the sources with true flux above the detection limit, therefore these source counts have to be corrected for effects such as completeness: confusion and Eddington bias. The correct counts will then mirror a true population rather than an observed population. The source count model we have created does have limitations, Franzen and Hopkins are only accurate at their own frequency and flux ranges, extrapolation beyond these limits will produce results that may not be in agreement with other measurements. However ranges within these limits are expected to agree well.

3.4 Source Structure

When looking at radio sources, the emission received from a source can be categorised into two different types, compact and extended emission. However it should be noted that reality is not black or white as sources can have compact components embedded in extended emission or multiple compact components spread throughout the source. Typically compact components are AGN core or hotspots within lobes whereas extended emission are the lobes of a radio galaxy. Using Interplanetary Scintillation [Chhetri et al. \(2018a\)](#) found that the scintillation strength was tied to structure of the source, with extended emission diluting

strong scintillation from compact sites at low resolution. This complex source structure complicates our simulation as we have to account for not just a simple extended or compact source but also how much of a given source’s flux may be due to compact sites or extended emission. For now we shall assume a simple duality between compact and extended sources.

While the definition of what is a compact or extended source is generally defined,

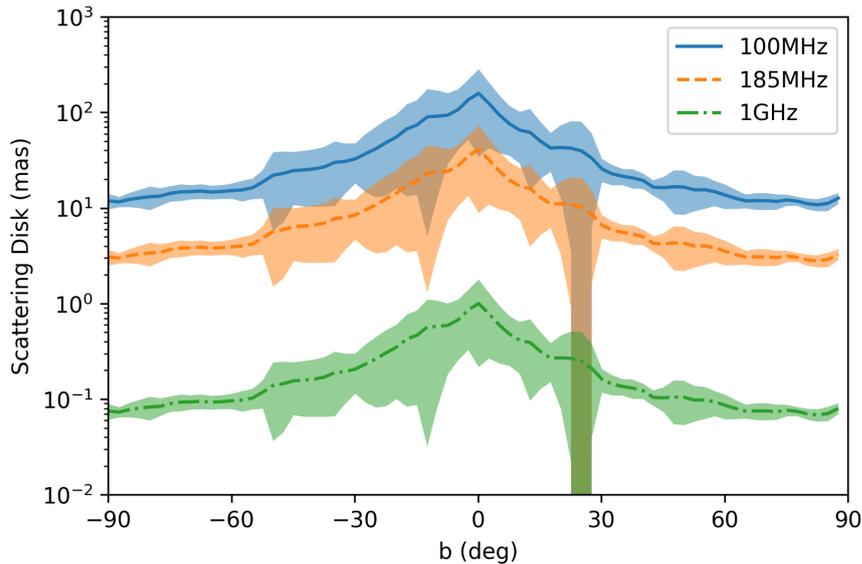


Figure 3.3: Figure showing the size scattering disk size (in milli-arcseconds) at different frequencies and latitudes. Data shows, 100 MHz (blue), 185 MHz (orange) and 1 GHz (green). The solid line represents the average value and the shaded cells are the standard deviation. Figure adapted from [Hancock et al. \(2019\)](#).

in terms of scattering there is a much more constraining definition. A source is compact if its angular size is smaller than that of the angular scattering scale which depends on the medium you are looking through. If a source is bigger than this scale it is extended and its scintillation is down weighted by the the magnitude of its extent. As we observe different regions of the sky at different frequencies, the critical angular scale at which a source scintillates will change, and therefore there will be no singularly defined compact size. This can be further seen in [Figure 3.3](#), where at different frequencies and latitudes the critical scattering scale changes. In our simulation we want a simple relation for the source size, and how many sources are compact or extended. We shall first find the ratio of compact to extended sources before determining the source size of

each fraction.

3.4.1 Source Structure and Size Distribution

We know that sources are complex and can involve several components embedded in extended emission, each producing a percentage of the flux, therefore it is hard to use a single number to describe them. Never the less we will try and find a ratio that can split the population of extra-galactic radio sources into compact or extended populations.

Spectral Index

Spectral Energy Distributions (SEDs) are intrinsically linked to the behaviour and type of source, we can use this to split the source population to extended or compact. SEDs can be approximated by the power law seen in [Equation 3.8](#).

$$S_\nu = \nu^\alpha \quad (3.8)$$

The synchrotron radiation emitted from extra-galactic sources can be broken down into optically thin (no absorption) or optically thick (some to all radiation absorbed). Extended regions are typically optically thin and have a spectral index of $\alpha \approx -0.7$, they have steep SEDs. Whereas compact regions such as AGN cores have a superposition of multiple turnover spectra which results in a flattening of the radio spectrum and therefore flat SEDs. We can put a limit on this by saying flat SEDs have a alpha value between -0.6 to 0.6 ($0.6 > \alpha > -0.6$) and steep SEDs anything less than -0.8 ($\alpha < -0.8$). To determine the fraction of flat and steep SEDs in the population we used the GLEAM catalogue ([Hurley-Walker et al., 2016](#)) and labeled sources with flat SED's as compact, whereas sources with steep SEDs are assumed to be extended. This fraction of extended to compact sources, according to this method, turned out to be $\sim 68\% : 32\%$ extended to compact. A benefit of this method is that it does not require us to resolve the sources, however this method does not account for mixed structures where, for example, a compact flat spectrum hot spot is embedded in an extended steep spectrum radio lobe.

Flux Ratio

We can use the ratio of peak flux to integrated flux to determine extended and point sources in their survey, this method comes from [Franzen et al. \(2019\)](#). Where the peak flux is the brightest pixel and the integrated flux is the flux of the entire source, for a point source these two values are the same, however, for an extended source the integrated flux will be larger than the peak flux. The method used is as follows: Define a parameter R , the log ratio of the total flux to the peak flux.

$$R = \ln(S/S_{\text{peak}}) \quad (3.9)$$

Where S is the total flux and S_{peak} is the peak flux. We can assume that the distribution is a Gaussian centered on zero with given RMS:

$$\sigma_R = \sqrt{\left(\frac{\sigma_S}{S}\right)^2 + \left(\frac{\sigma_{S_{\text{peak}}}}{S_{\text{peak}}}\right)^2} \quad (3.10)$$

Where σ_R is the uncertainty in R , σ_S is the uncertainty in the total flux and $\sigma_{S_{\text{peak}}}$ is the uncertainty in the peak flux. Therefore to detect a extended source at the 2 sigma level, we must have $R > 2\sigma_R$:

$$R > 2\sqrt{\left(\frac{\sigma_S}{S}\right)^2 + \left(\frac{\sigma_{S_{\text{peak}}}}{S_{\text{peak}}}\right)^2} \quad (3.11)$$

We used this method to search several catalogues of varying frequency and angular resolution to work out the percentage of sources that are found to be extended. The catalogues were initially selected if they had both a integrated and peak flux reported in their catalogues, further the initial surveys were centred around 151 MHz to compared how the extended percentage changed with other factors such as sensitivity and angular resolution. We then selected several surveys of increasing frequency to see how this fractional percentage increased with frequency. The surveys selected will now be described in summary.

TIFR GMRT Sky Survey (TGSS) - [Intema et al. \(2017\)](#), is a 150 MHz catalogue using the Giant Metrewave Radio Telescope (GMRT), which is an array of 30x45m telescope dishes. The catalogue covers a 99.5% of the sky north of -53 degrees declination, with a resolution of 25 arc seconds and a sensitivity of 3.5

milli-Jansky per beam.

GaLactic and Extragalactic All-sky Murchison Widefield Array (GLEAM) - [Hurley-Walker et al. \(2016\)](#), is a 74-231 MHz that use the MWA telescope located in Western Australia. The survey covered 24,831 square degrees and covered declinations south of +30 degrees, but didnt include sources within 10 degrees of the galactic plane. GLEAM also has a resolution of approximately 100 arc seconds and a sensitivity of 8 milli-Jansky per beam.

The LOFAR Two-metre Sky Survey (LoTSS) - [Shimwell et al. \(2019\)](#), is a deep survey (0.1 milli-Jansky per beam) and high resolution survey (6 arc-seconds) that observes at a range of frequencies between 120-168 MHz, it is still being completed but aims to map the entire northern sky. As of 2019 it has completed 20% of its observations.

Sydney University Molonglo Sky Survey (SUMSS) - [Mauch et al. \(2003\)](#), is a 843 MHz survey using the Molonglo Observatory Synthesis Telescope (MOST) which operates from New South Wales using two very long (778mx12m) cylindrical paraboloids. This survey has an angular resolution of 45 arc seconds and a sensitivity of 1 milli-Jansky per beam and it observed all of the sky south of -30 degrees.

Survey Catalogue	Frequency (MHz)	Sensitivity (mJy/beam)	Angular Res (arcsec)	Number of Sources	Extended Sources	Extended Percentage
TGSS	150	2-5	25	623,096	147,624	24%
GLEAM	151	6-10	100	259,059	8,804	3.4%
LOTSS	120-168	0.1	6	44,172	5,493	12%
SUMSS	843	1	45	193,360	39,298	20%

Table 3.2: Table showing the statistics of each survey and then the calculations done from the catalogues.

From [Table 3.2](#) there is no clear trend in the percentage of sources that are

extended based on frequency. This is due to the multitude of other changing parameters in each survey: sensitivity, angular resolution, population. However as the angular resolution of beam increases we would expect this method to increase in accuracy, as we better resolution would be able to distinguish between compact and extended sources. As an example GLEAM has low angular resolution and therefore the majority of the sources it observes are compact compared with the synthesised beam. Whereas, in TGSS, which has 4 times the resolution, and roughly the same operating frequency and sensitivity, we see 8 times the number of extended sources. In summary this method only determines if sources are smaller or larger (extended) than the resolution, so if we wanted to get milli-arcsecond resolution then we would have to use VLBI, which would take exorbitant amount of data and time to survey some significant fraction of the sky, and so we are left with these large scale surveys. Due to the properties changing between each telescope it would be hard to chose a telescope as a base for the entire population, if instead we could remove the limitation of using a telescope and instead use the intervening medium to determine source size, that would remove this limitation.

IPS statistics

Using Interplanetary Scintillation (IPS) and the MWA, [Chhetri et al. \(2018a\)](#) investigated scintillating sources using the Interplanetary Medium (IPM) as a telescope. Due to the typical lengthscale of inhomogeneties within the IPM at the time of observation, the critical angular scale can be calculated and any sources observed to scintillate must be smaller than this size. From this [Chhetri et al. \(2018a\)](#) were able to identify sub-arcsecond components across 900 deg² area of the sky, and grouped their high signal to noise sources into three categories: strong scintillators (8.9%), moderate scintillators (23.4%), weak/non-scintillators (53.4%) and undefined (14.3%). They further investigated these sources by comparing them with a higher resolution survey, the Faint Images of the Radio Sky at Twenty-Centimeters (FIRST) [Becker et al. \(1995\)](#). The FIRST survey has a resolution of 5 arc-seconds and a sensitivity of 0.15 mJy and was created using observations from the Very Large Array (VLA) at 1.4 GHz, the VLA is a Y shaped array of 28 25 metre dishes located in New Mexico, USA. From this com-

parison they found that strong scintillators usually had one component whereas moderate scintillators had a single or dual compact components embedded in extended emission and weak scintillators were dominated by extended emission or had multiple components. Additionally they calculate the normalised SED's for each group as seen in [Table 3.3](#), this is done by cross-matching their sources with both the GLEAM and NVSS.

Strength	α_{76}^{277}	N	α_{162}^{1400}	N
Strong	-0.37	37	-0.68	37
Moderate	-0.7	97	-0.77	91
Weak/Non	-0.87	221	-0.81	171
Undefined		59		49
Total	-0.82	414	-0.8	347

Table 3.3: IPS SED's for different strengths of scintillation. Center column is cross-matched using GLEAM and right column data using NVSS. Data taken from [Chhetri et al. \(2018a\)](#).

If we assume that the IPS sample of [Chhetri et al. \(2018a\)](#) is representative of the entire population we could assume that the strong and moderate (37+91) scintillators represent our scintillating population, which gives a fraction of 128/347 (37%). Further, [Chhetri et al. \(2018b\)](#) finds 93/247 (38%) of their sources are point sources (≤ 0.3 arcsec) in their source count catalogue, which is in rough agreement with the previous figure for expected scintillators. It is important to note that the critical angular scale for angular scattering at the IPS regime is an order of magnitude larger than values typical for the ISS regime. This would mean that the observed compact source population will be larger than the population small enough to exhibit RISS and therefore should be thought of as an upper limit to the number of sources that could be scintillating in the ISS regime.

Final Distribution Model

We have discussed three different models, SED ratio, flux ratio, and IPS statistics. The SED and flux ratios are dependent on the catalogue you use and

the telescope, therefore we are limited to the frequency, resolution, and sensitivity of the telescope. However this can be limited by using catalogues that have large frequency ranges such as GLEAM or multiple different frequency catalogues. With the IPS statistic there is no telescope dependence, while [Chhetri et al. \(2018a\)](#) use the MWA telescope to observe IPS, they are actually using the IPM as the telescope to determine source sizes, and so are not beholden to the same limitations that the first two methods are. However, we can see that the SED ratio (32%) is consistent with that measured by the IPS statistics (37%). For this reason we shall implement the IPS statistics as an upper limit on the number of compact source within our model: therefore we assume that 37% of our sources are small enough (compact) to scintillate to a maximum degree.

Source Size

Now that we have the the ratio of compact and extended sources we need to incorporate the source size of each fraction. The ideal method to determine source size would be to use VLBI and determine the size a structure of all the sources, however this would take an egregious amount of time to complete as well as an unimaginable amount of data. Therefore we will instead use simplified models and assumptions.

Linear Size and Distances

The most simplistic method would be to assume the typical redshift and linear size of a compact source and extended source. The typical redshift for a AGN is $z = 1 - 2$ ([Williams et al., 2018](#)) which gives a distance of 1.774 Gpc, if the compact site of the AGN is approximately 1 pc, this would return an angular size of 0.1 milli-arcseconds. Whereas an extended source would have a linear size significantly larger than the compact AGN core, a typical galaxy such as the Milky Way are approximately 30 kpc wide. Using the same redshift, returns approximate angular size of 5 arcseconds.

Population	Fraction	Size
Compact	37%	0.1 mas
Extended	63%	Equation 3.12

Table 3.4: Table showing the final fraction and size of each population: compact and extended.

Windhorst Model

The previous linear model is over simplistic and does not take into account any other factors of the sources instead we can use a observation data based model for these sources. [Windhorst et al. \(1990\)](#) provides a model for expected source size at given flux ([Equation 3.12](#)). Furthermore ([Williams et al., 2016](#)) extrapolate this model from 1400 MHz down to 150 MHz and find that it is in agreement with LOFAR observations at 150 MHz , showing that it can be used at multiple frequencies.

$$\theta_w = 2'' S_{1.4\text{GHz}}^{0.3} \quad (3.12)$$

Where θ_w is the source size in arc seconds, S is the flux of the source at 1.4 GHz in mJy. For example a distribution of fluxes over 1mJy to 1Jy at 185MHz would have source size range of 1.23'' to 9.77''. We will apply this model to our extended sources and the flux generated by the source counts.

Final Size Model

For the final model used in the simulation, for extended sources it was decided that the [Windhorst et al. \(1990\)](#) model would be used based on their flux. For the compact sources, as we just require a small enough size so that it is scintillating we used the linear model for this (0.1 milli-arcseconds).

3.5 Variability Calculation

We now have a program that generates positions, flux and source size. Now this can be put into RISS19 to return variability parameters. Specifically the modu-

lation index can be used determine the variability of our sources. We can get the ‘true’ variability from RISS19, however because we are trying to simulate *observed* variability, we will need to apply further selection criteria in order to mimic what a telescope would observe. To do this we make two selection cuts: magnitude and significance of variability. The first selection cut finds sources above the modulation cutoff ($m_c = 5\%$) this assess the magnitude of the variability.

$$m \geq m_c \quad (3.13)$$

Additionally we do a second selection because we are trying to mimic this *observed* variability, and we know that telescopes have some measurement uncertainty and noise associated with them. An observer using a telescope would only select variable sources that exhibited modulation significantly higher than the rms noise of the image from the telescope. The second cut determines the significance of the variability by selecting sources that have flux modulation five times greater than the rms noise of the image.

$$mS_\nu \geq 5\sigma_{\text{rms}} \quad (3.14)$$

We now have two selection cuts, one for the magnitude of variability and a second for the significance of the variability when compared to some associated noise of the telescope. Now that we have made these selections we can collate our results together to find the surface density.

Simulation Results

After generating source positions, flux, structure and size, we passed these variables through RISS19 which returned variability parameters, these parameters allowed us to make selections cuts to determine how many variable sources we have found. We can now calculate the surface density (ρ), the surface density, sometimes referred to as the areal sky density is the the number of variable (n_v) sources per degree squared (Ω) (Equation 3.15).

$$\rho = \frac{n_v}{\Omega} \quad (3.15)$$

However if we detect less than one variable source, we still want to be able to say something meaningful about the area of the sky. We can instead place an upper limit on the surface density for this area. [Bell et al. \(2011\)](#) (Equation 1; Equation 3.16) do this by using Poisson statistics to place a confidence limit on what we should expect from an area of the sky.

$$P(n) = \exp(-\rho\Omega) \quad (3.16)$$

$$\rho = \frac{\ln(P(n))}{-\Omega} \quad (3.17)$$

For example, in the simulation we will apply a 95% confidence limit on the surface density, to do this we set $P(n) = 0.05$ and re-arrange for ρ , and input the area.

3.6 The Simulation

We now have all the individual parts for the simulation, we will now combine them together into one program. [Figure 3.4](#) shows a flowchart of the simulation program including the main inputs and simulation parameters used. The simulation can take in multiple parameters from the user: lower flux limit, upper flux limit, modulation cutoff, observation time, frequency, number of iterations, region file name and output file name. Most of these have default values but the region file name is required and, the simulation will output key details to the user terminal unless output file name is specified. The output file is generated in two parts, the first part is the data file which contains the mean and standard deviation for each iteration's $H\alpha$, modulation, timescale, scattering disk size and just the calculation for the surface density. The second output file is the results file which contains the mean and standard deviation for all of the iterations, as well as the input parameters into the simulation. We can summarise the simulation as follows:

1. Program reads in command line input parameters: lower and upper flux limits, timescale of observation, frequency, modulation cutoff, number of iterations to run for, region file, and output filename
2. Generates source counts for 3 times the lower flux limit.
3. Generates uniformly distributed positions within the input area of the sky.

4. Assigns the source type as either compact (37%) or extended (63%) based on the fraction given by IPS statistics.
5. Each source type has a size assigned to it, the compact fraction is assigned a value of 0.02 milli-arcseconds, and the extended fraction uses the flux in combination with the Windhorst model to generate its size (Equation 3.12).
6. The source size and position of each source is then fed into RISS19 which returns variability parameters: timescale of modulation, critical angular scale, and modulation for the observed timescale.
7. These variability parameters are used in combination with the flux to determine significant variables from the two selections cuts made for magnitude of variability (Equation 3.13) and significance of variability compared to image noise (Equation 3.14).
8. The number of significant variables are then used with the area to calculate the surface density (also called areal sky density) (Equation 3.15), if less than one source is detected an upper limit on the surface density is implemented using Poisson statistics (Equation 3.16).
9. The program then repeats this process for some number of iterations, once it has done this it returns a results file containing the mean and standard deviation values for the whole run.

We now have a simulation that can generate sources and their properties. We have then combined this with RISS19 which allows us to predict population trends for given regions of the sky. We now want to see how our simulation compares to surveys from the literature, and determine any limitations to the simulation. To do this we have selected a variety of surveys with ranging, frequency, flux density, area and timescale. The following surveys were selected for several reasons: i) the survey reported a variability cutoff parameter that could be converted into a modulation cutoff, and, ii) the region of sky was relatively easy to create by hand using MIMAS from AegeanTools (Hancock et al., 2018). I will now give a short description for each of the nine surveys selected.

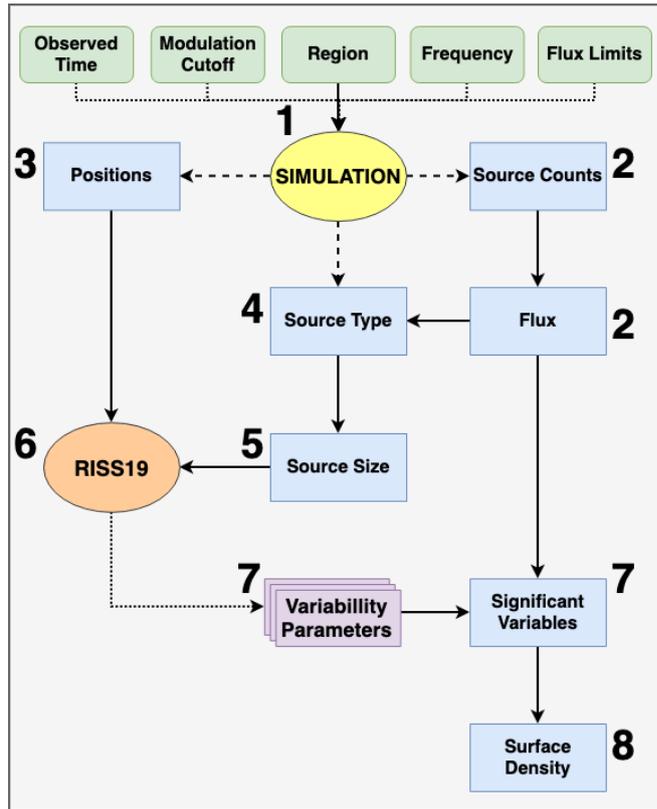


Figure 3.4: The simulation program flowchart, where green boxes are the inputs from the user, blue boxes are the simulation outputs, and the purple box is the RISS19 output.

HGL - [Hancock et al. \(2019\)](#), this is a 185 MHz survey using the MWA, designed to test the RISS19 model and how variability changes as a function of latitude. It observed 4 circular fields ranging in latitude from -16 to -90 degrees (galactic coordinates) with a radius of 15 degrees. There is no currently published paper for this survey, however Dr. Paul Hancock kindly provided results for use in this project. They found a surface density of $\rho = 0.0237\text{deg}^{-2}$ on timescales of 7 to 455 days.

TH19 - [Tingay & Hancock \(2019\)](#), is a survey observing the Kepler Fields 3,4, and 5 at 154/185 MHz using the MWA, the aim of this survey was to search for low frequency variability and transients. We will use only the 185 MHz data for a direct comparison to the HGL survey. To select variables they implement a de-biased modulation cutoff of 5% and a probability that the light-curve could be produced due to random noise to be $P \leq 0.001$, from this they did not find

any variables and instead implemented an upper limit of $\rho \leq 0.0002\text{deg}^{-2}$ on timescales of 2 to 20 days.

M13 - (Mooley et al., 2013), is a 49 epoch survey that uses archival data of the Extended Chandra Deep Field South region at 1.4 GHz using the Very Large Array, with the goal to find transients and variables at sub-milliJansky levels. They select variables using two criteria i) a confidence cut at 4σ using chi squared probability, ii) a modulation index greater than 50% for ‘strong’ variables. They do not detect any ‘strong’ variables however they do detect some ‘significant’ variables. From this they find a surface density of $\rho = 1.17\text{deg}^{-2}$ on timescales of 7 to 84 days. However as they do not define the modulation cutoff for ‘significant’ variables we will assume this to be less than 50% so we will select 30%.

PDS - (Hancock et al., 2016) is a paper that studies radio variability in the archival data of the 1.4 GHz Phoenix Deep Survey (Hodge et al., 2013), the aim of this survey is to compare variability found in this data to other similar surveys, as well as resolve discrepancies between other surveys findings. They select variables using a de-biased modulation cutoff of 5%, and a significance of 3σ variability from the light curve. From this they find a surface density of $\rho = 0.98\text{deg}^{-2}$ on timescales of 6 months to 8 years.

F94 - Frail et al. (1994), is a 0.4/1.4 GHz survey looking for a radio counterpart to a gamma-ray burst on 1st March 1994. They use the Dominion Radio Astrophysical Observatory Synthesis Telescope, which is a telescope located in British Columbia Canada and consist of seven nine metre antennas with a 600 metre baseline in the east-west directionm this telescope is capable of observing at both 400 MHz and 1.4 GHz. F94 used this telescope to obtain 15 images at a 0.7 mJy level at 1.4 GHz. We will only use the 1.4 GHz data for direct comparison to our other 1.4 GHz surveys, as the 400 MHz results are not described fully within the paper. The criteria for variables was that they had to be 5σ (3.5 mJy) above the daily noise level. However no variables were found and they put a 3.5

mJy limit on transients in the image. We converted this into a surface density limit using Equation 3.16 which returns $\rho \leq 0.111\text{deg}^{-2}$ on timescales of 3 to 99 days.

H13 - (Hodge et al., 2013), is a blind survey that compares two archival epochs of the FIRST survey with new data from the SDSS stripe 82 at 1.4 GHz, the overlapping region has an area of 60 deg². The aim of this survey is to find variables within this region, they do this by using the following criteria: i) Sources are 5σ , ii) the SDSS data must be a point source, iii) the SDSS data must be brighter than the FIRST and iv) variability amplitude greater than 30%. From this they find 89 variable sources for a surface density of $\rho = 1.48\text{deg}^{-2}$ on timescales of 7 to 14 years.

T11 - (Thyagarajan et al., 2011), is a large scale 1.4 GHz survey comprising of $\sim 55,000$ snapshot images and covering an area of 8444 deg² from the FIRST survey. The aim of this survey was to be able to have a large enough source population to match to multi-wavelength catalogues to further study the variable objects. They select sources that have a variability 5 standard deviations from the mean, using a chi squared probability distribution. However, as this is not easily implementable into our model, we selected a 20% modulation cutoff for this survey. From this they find 1627 variable sources for a surface density of $\rho = 0.582\text{deg}^{-2}$ on timescales 1hr to 2 years.

B15 - (Bell et al., 2015), is a 3 epoch survey that observes the Extended Chandra Deep Field South region at 5.5 GHz with the Australian Telescope Compact Array (ATCA), which is comprised of six 22 metre antennas, in New South Wales. This survey was designed to be the deepest search for variables and transients at 5.5 GHz. They select variables based on the chi squared probability of $P \leq 0.001$ which found 7 variable sources, they then used the de-biased modulation index to remove any low modulation sources. This left 4 variables sources for a surface density of $\rho = 3.23\text{deg}^{-2}$ on timescales 18 to 913 days. The selection criteria does not have any defined modulation cut off, however the lowest de-biased modula-

tion index of the 4 variables was 2% and so we will use this as our cutoff.

The summary of parameters for each survey can be seen in [Table 3.5](#), and the results for each in [Table 3.7](#), [Figure 3.8](#). While these surveys are some of the best that could be found, there are still some typical problems found across each survey. These problems include the different styles of reporting and classification of data in transient and variable science. Further as we do not have access to most of the original data, the source selection done has to be inferred from the methods within the paper, here there will lie some discrepancy, as if a paper removes a certain selection of sources, our simulation will not account for that. The surveys which pose the least amount of issues which I expect to have the greatest agreement would be HGL, TH19, and the PDS survey as these three surveys use consistent and simple methods for determining the variability of their sources.

Survey Names	Frequency (MHz)	Area (deg ²)	N_S (N/deg ²)	Timescale	Flux Density (mJy)	Modulation Cutoff
HGL	185	2442.29	7.47	7 - 455d	17	0.05
TH19	185	2775.18	2.30	2 - 20d	70	0.05
M13	1400	0.3211	2086	7 - 84d	0.0074	0.3
PDS	1400	5.97	31.8	0.5 - 8.2yr	1.4	0.05
F94	1400	9	51.95	3 - 99d	0.7	0.2
H13	1400	60	40.21	7 - 14yr	1	0.3
T11	1400	8444	40.24	1hr - 2yr	1	0.2
B15	5500	0.3211	2122	18 - 913d	0.021	0.02

Table 3.5: The parameters of each survey: The survey names, frequency in MHz, the area in degrees squared, the number of sources per degree squared, the timescale of modulation (minimum-maximum), the lower flux density of the survey in milli-Jankys and the modulation cutoff parameter used.

3.7 Results and Discussion

We now have a selection of surveys that we can use to compare with our simulated versions of these surveys. Firstly we will look at the HGL survey as this

is the survey we expect the highest agreement from as well as having the catalogue for this survey, then we will compare to a large selection of surveys. In this section we will refer to two different simulation strategies - point and ratio. This just determines whether the population will be full of point sources (100% point strategy) or a ratio of compact to extended sources (37% point strategy). We implement these strategies as we expect a population of point (compact) sources to modulate maximally, and therefore provide an upper limit to the variability, whereas a mixed population of compact and extended objects will have modulation closer to a true population. In terms of code it changes the ratio of compact to extended sources so that we have two populations, one consisting of 100% compact point sources and another of 37%.

3.7.1 Comparison to HGL

Using the simulation program we can determine population statistics for the HGL region, we have separated this region from the others as we have the survey catalogue for this survey and can perform further tests and comparisons using this information. For this survey we used the maximum timescale of 448 days, a modulation cutoff of 5%, a lower flux limit of 17 mJy, upper flux limit of 10 Jy, as well as the region files provided by Dr. Paul Hancock. Before running the simulation we must first understand that when observing with a radio telescope there are associated sensitivity limits and noise of the telescope. These properties mean that we have to restrain what sources we select to be real or false, this is due to the fact that as we approach lower flux there is a higher chance for false detections to be made and so a detection threshold is set to minimise false detections while maximising the detection of real faint sources. This leads to the preferential detection of brighter sources, for example most telescopes have some rms sensitivity and select sources that are greater than 5 standard deviations above that rms. Our simulation does not account for this bias it means it we will find more sources at lower fluxes than a real telescope would.

In [Figure 3.5](#) we can see the histograms for the number of sources versus flux

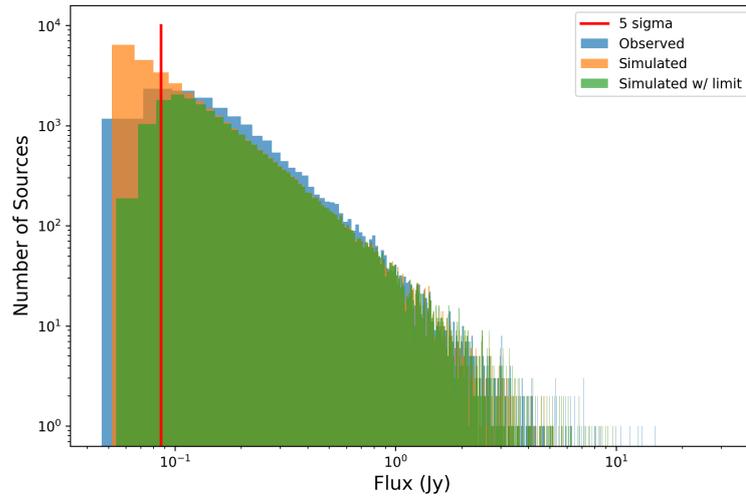


Figure 3.5: Figure showing the number of sources versus the flux found the HGL survey (blue), the simulation (orange) and the simulation with an estimated completeness limit implemented (green). The 5 sigma cutoff is also shown as a red line.

for HGL survey and its simulation counterpart. We can see at approximately 0.1 Jy close to the 5 sigma limit (85 mJy) there is a turnover in the observed number of sources (blue), whereas the simulation continues to increase. We can try partly resolve this by implementing a basic completeness limit cutoff (green). This reduces the number of sources per deg squared from 13.75 to 10.25, which reduces the error from 107% to 36% when compared to HGL’s 7.47 sources per degree squared. This basic completeness limit was implemented by assigning a 20% flux uncertainty to each source and using probability to select whether it was detected above the detection threshold or not.

We have found one reason for the over-prediction of sources per degree squared, another issue is that we have assumed the MWA rms sensitivity to be consistent throughout the image, however this is not true and it instead has a large distribution of rms changing across the image, and because the counts are large for fainter flux, a small change in the rms has a large change on the number of sources. For 3 sigma upper and lower limits the rms ranges from 9 mJy to 36 mJy with a mean of 17 mJy (the assumed sensitivity). If we instead used the 3 sigma limits as the rms sensitivity input for the simulation we would get

21.96 and 7.26 sources per degree squared respectively. We can use this to put an error bar on our counts: 14 ± 7 , which puts us in the range of the observed HGL counts (7.47). We can also run these two limits through the simulation to see how this changes the predicted surface density. In Table 3.6 we can see the results for each population (100%/37% point sources) for upper and lower limits on the rms sensitivity, where the upper/lower limits refer to smaller/larger rms sensitivity values respectively. We can see that as expected the surface density scales proportionally to the number of sources per degree squared. We can use the upper and lower limits to put a new uncertainty on the mean result by taking half the difference: HGL:100% Point $\rho = 1 \pm 1 \text{ deg}^{-2}$ and HGL:37% Point $\rho = 0.5 \pm 0.4 \text{ deg}^{-2}$.

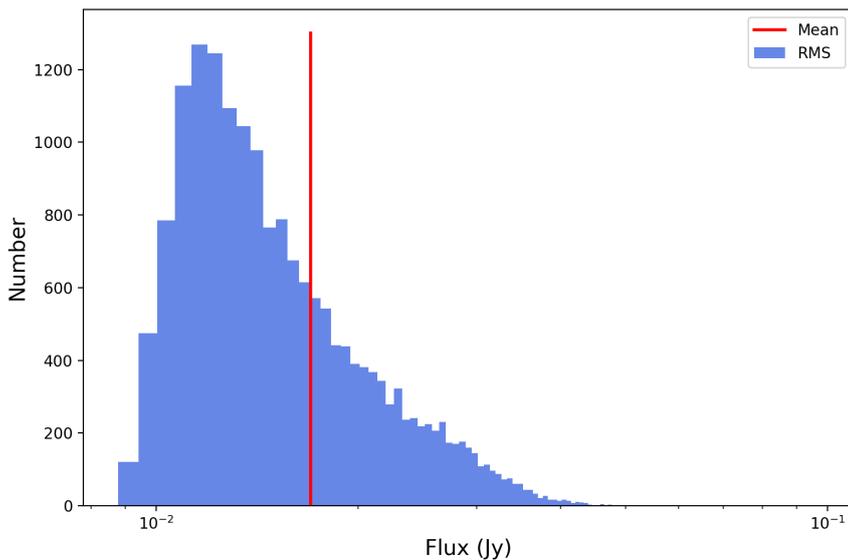


Figure 3.6: The flux in Jansky of the local rms found in the HGL survey.

Even though the surface density prediction has limitations, we can still use the data from the simulation, as we can instead look at the bounds on the variability that the simulation predicts. In Figure 3.7 we have taken the $H\alpha$ and modulation for each point above the 5% cutoff and created two lines of best fits for each population: 100% and 36% point sources. We have then plotted a shaded area where the upper bound is the 100% point population and the lower bound is the 37% population. We can see that the trend predicted by the simulation

Name	N_S (N/deg ²)	Variable Sources (N %)	ρ (deg ⁻²)
HGL	7.47	58(0.318)	0.001
100% Point-3 σ	21.96	6550(12.21)	2.68 ± 0.02
100% Point: Mean	13.75	3196(9.52)	1.31 ± 0.02
100% Point+3 σ	7.26	1138(6.42)	0.47 ± 0.01
37% Point-3 σ	21.96	2404(4.48)	0.98 ± 0.02
37% Point: Mean	13.75	1178(3.51)	0.48 ± 0.01
37% Point+3 σ	7.26	417(2.35)	0.171 ± 0.008

Table 3.6: The results for the simulation for 100% and 37% point source populations when assuming the rms mean and plus/minus three sigma.

program agrees with the majority of the points from the HGL survey. This shows that while the simulation may over predict the number of variables, the bounds it predicts is still in agreement with results, furthermore this bounding function can be used to argue for or against ISS as the source of variability within a survey or to inform plans for a future survey.

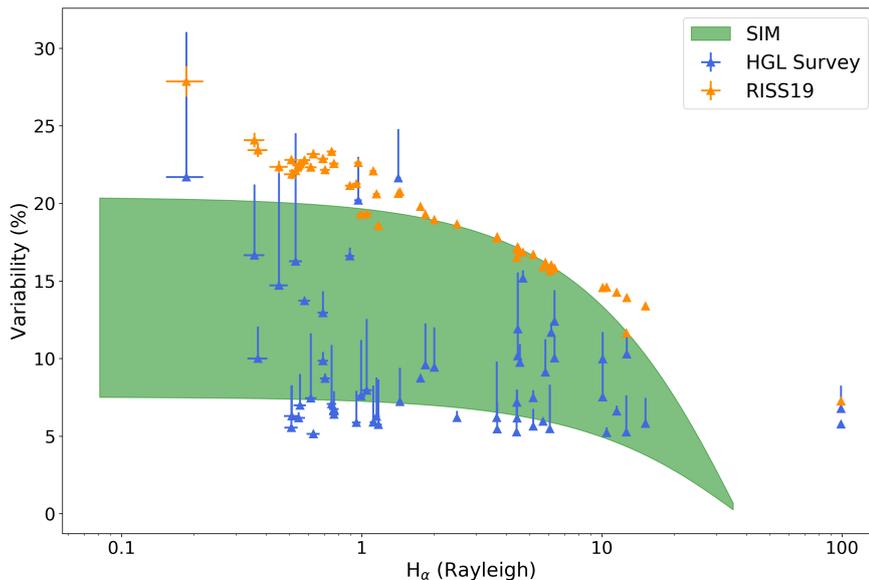


Figure 3.7: The variability versus $H\alpha$ intensity for the HGL survey. The data shown: HGL Survey (blue circles), RISS19 model at its max modulation (orange triangles), and the green region simulated for point sources (upper limit) and 37% Point sources (lower limit). Figure adapted from [Hancock et al. \(2019\)](#).

3.7.2 Comparison to Surveys

We have compared our simulated observation to the HGL survey and found a good qualitative agreement between the two, but note that there are limitations when converting between the ‘true’ and observed variability of the two populations. We will now compare our simulated results to the surveys listed in [Table 3.5](#). For this comparison we have used the maximum timescale for each survey. Results for each population (100%/37% point sources) are shown in [Table 3.7](#) and [Figure 3.8](#).

Looking at the results in [Table 3.7](#) there is a common trend in both the number of variable sources and the surface density: the ratio of variable sources to non variable sources (37% to 63%) matches our input fraction of point sources almost exactly, with some variance due to the source sizes and how they are distributed. This is because the simulation is generating source sizes which are much smaller or larger than the critical scattering angle. The RISS19 model determines

Survey	N_S (N/deg ²)	Variable Sources (N %)	ρ (deg ⁻²)
HGL	7.47	58(0.318)	0.001
HGL:100% Point	14 \pm 7	3196(9.52)	1 \pm 1
HGL:37% Point	14 \pm 7	1178(3.51)	0.5 \pm 0.4
TH19	2.3	0(0.0)	0.0
TH19:100% Point	3.95	0(0.0)	\leq 0.001
TH19:37% Point	3.95	0(0.0)	\leq 0.001
M13	1865.4	7(1.169)	0.454
M13:100% Point	2085.7	309(21.79)	454 \pm 12
M13:37% Point	2085.46	114(8.08)	168 \pm 13
PDS	438.69	8(0.305)	0.769
PDS:100% Point	31.68	148(52.36)	17 \pm 1
PDS:37% Point	31.66	54(19.2)	6.1 \pm 0.9
F94	27.22	0(0.0)	\leq 0.111
F94:100% Point	51.95	213(44.29)	23 \pm 1
F94:37% Point	51.76	77(16.23)	8.4 \pm 0.9
H13	23.93	89(6.198)	0.742
H13:100% Point	40.21	1351(50.05)	20.1 \pm 0.3
H13:37% Point	40.26	498(18.44)	7.4 \pm 0.3
T11	33.09	1627(0.582)	0.193
T11:100% Point	40.24	177303(51.74)	20.82 \pm 0.01
T11:37% Point	40.24	65422(19.09)	7.68 \pm 0.02
B15	386.16	4(3.226)	12.457
B15:100% Point	2121.78	45(3.13)	66 \pm 8
B15:37% Point	2121.59	16(1.17)	25 \pm 6

Table 3.7: The simulation results for each survey, assuming 100% and a 37% point source population. Showing number of sources per degree squared (N_S) the total number of variable sources and the percentage of the population and the surface density which is the number of variable sources per degree square (ρ).

the variability of the source based on the source size and critical scattering angle for that direction, and because the source sizes are either magnitudes smaller or

larger than this angle we are generating a bimodal population of variable or non-variable sources. Due to this, we don't actually need to re-run the simulation to determine the number of variable sources as this will be determined by the fraction of point sources.

Survey	Frequency (MHz)	Scattering Angle (mas)	Point Sources (%)
TH19	185	8.3	–
HGL	185	16.3	0.09
F94	1400	0.06	0.55
T11	1400	0.049	1.92
M13	1400	0.075	0.17
H13	1400	0.048	6.23
PDS	1400	0.037	8.67
Average	1400	0.053	3.51
B15	5500	0.0037	19.61

Table 3.8: For each frequency band the critical angular (mas) scale and fractional percentage of point sources (%) required to bring the simulation results into line with the survey observed results. We can then say that for our source distribution model that X% of sources must be smaller than the given critical scattering angle.

For any single survey we could adjust the fraction of point sources so that our predictions match the surveys, we would then be using the simulation to determine the number of sources smaller than the scattering disk to match this surveys results. In [Table 3.8](#) we have separated the surveys into there frequency bands and displayed their critical scattering angle in milli-arcseconds, as well as the fraction of point sources that would be required to bring our simulation results into line with the observed results of the surveys. For each frequency range we get the following point source populations seen in [Figure 3.8](#): 0.09% at 185 MHz, 3.51% for the 1400 MHz surveys, and 19.61% for the 5500 MHz survey. For the 1400 MHz range we have one notable outlier the M13 survey, this survey

is significantly deeper than the other surveys with a minimum flux density of 74 micro-Janskys. At lower flux densities there are increasingly more extended SFG-like sources (see Figure 1.5a) and so our simulation is overpredicting the number of compact scintillating sources.

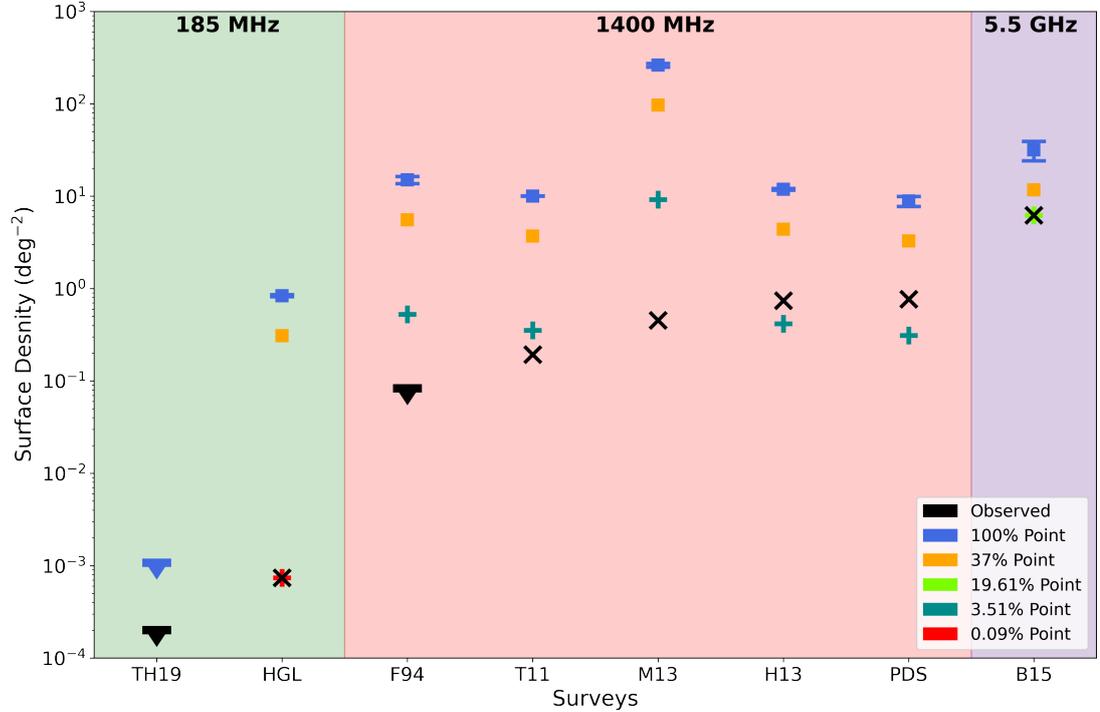


Figure 3.8: Surface density for each survey (black crosses), the simulations for 100% point source population (blue squares), 37% population (orange squares) as seen in Table 3.7. Bars with triangles indicate upper limits where less than one variable source was detected within the survey. The plot has been separated into three frequencies, 185 MHz (green), 1400 MHz (red), and 5.5 GHz (purple). For each frequency the factor of point sources required to match the target survey has been plotted as: 19.61% for 5.5 GHz (lime plus), for 1400 MHz 3.51% (teal plus), and for 185 MHz 0.09% (red plus).

Compact sources such as AGN cores and hot-spots tend to have a flat spectrum, being equally bright at all frequencies. Extended sources such as the lobes of radio galaxies, or the disk of star forming galaxies, tend to have an inverted spectrum, being brighter at low frequencies and fainter at high frequencies. All up this means that we would expect to see mostly extended sources in a low frequency survey, and mostly compact sources in a high frequency survey. This is in line with what we see in Table 3.8 - fewer compact sources at 185 MHz,

and more at 5.5 GHz. This means that the underlying model of our simulation is indeed correct and follows the trend as expected, however the source count function and source size distribution used is insufficient to describe the low frequency compact source population. One notable results is that of B15, where the simulation predicts that 19.61% of sources must be smaller than 3.7 micro-arcseconds. This source size restriction is stricter than one would expect, on the timescales that the B15 survey observed at, it is likely that some fraction of the variability observed is indeed intrinsic and not extrinsic. Further, at 5.5 GHz we are entering the transitional regime between strong and weak scattering as we increase in frequency (Figure 1.12).

3.7.3 Current State

In Chapter 2 we addressed issues and updated the RISS19 model to demonstrate its ability to put limits on variability due to RISS for observations. In this chapter we have attempted to use the RISS19 model to predict variability for surveys by generating sources within that area of the sky and running them through RISS19. While we have obtained results that demonstrate the underlying functionality of RISS19 issues have been highlighted regarding the data set and models used for generating compact source populations.

From simulating other surveys we found that in Table 3.7, the source size distribution determines the fraction of variable sources almost entirely. Therefore this component of the simulation is required to be accurate if we wish to produce accurate results. Unfortunately there is a dearth of information regarding the source size distribution and the source counts of that distribution. While we did have issues simulating the surveys surface density accurately, we were still able to put limits on the percentage of the population that had to be smaller than the critical angular scale of scattering (Table 3.8).

We have made progress in the creation of a simulation which utilises the

RISS19 model to determine variability for a given region of the sky, however until more information becomes available for source sizes and their counts, our simulation will be severely limited by the distribution that we choose, and this will become future work when more of this information becomes available.

Chapter 4

Conclusion

In this thesis we explored the issue of radio variability at sub gigahertz frequencies, and concluded from the literature that the dominant cause of variability in this regime is extrinsic, specifically refractive interstellar scintillation due to the interstellar medium. With the deployment and construction of new large scale radio telescopes, a model that could predict these effects would be useful to an astronomer who wanted to determine the variability of their sources or survey region. To create a model of these effects, we then demonstrated how we can use tracers of the interstellar medium, such as hydrogen alpha as a proxy for turbulence and inhomogeneties, which can in turn be used to predict scintillation and scattering effects. The goal throughout this thesis has been to expand on the base SM2017 model so that it can be used for predicting variability for given line of sights, as well generating artificial sources to then simulate the population variability statistics for given regions of the sky, in attempt to model what an observer would see.

In [Chapter 2](#) we built upon the SM2017 model by adding usability for extended sources, and including uncertainty propagation. The updated SM2017 model became the RISS19 model. This model uses tracers of the ISM to predict variability for given line of sights provided by a user. This model was then used in the publication of [Hancock et al. \(2019\)](#), where the HGL survey was compared to predictions for RISS19, this paper described results and limitations of the model. In this thesis we addressed several of these limitations by obtaining new

H α data, as well as implementing new distance models. After these improvements were made, the new RISS19 model was tested on multi-frequency data using SED data provided by colleague Kathryn Ross. From these tests it was found that the RISS19 model can be given lines of sight, at different frequencies and return predictions that act as an upper limit for the extrinsic variability that one should expect to see from a source.

We then expanded upon this work in [Chapter 3](#) by building a simulation to generate extra-galactic radio sources within given regions of the sky. This was achieved by combining theory and observational results to generate source's position, flux and size. However, major limitations in doing this were noted as there are no accurate compact source counts at low radio frequency and therefore some amount of uncertainty lies within the source counts generated. We simulated 8 different surveys and found that the simulation generally over-predicts the fraction and surface density of variable sources.. Using information obtained from the source catalogue of the HGL survey we were more closely able to analyse these results, we found that discrepancies occurred due to the completeness limit and the changing rms sensitivity of the telescopes beam not being accounted for by the simulation. Furthermore we speculated that due the MWA's large angular resolution, there must be dilution of compact components by extended emission therefore decreasing the surface density seen by the survey. After noting these effects we then compared the simulations to 8 other surveys. From these results we determined that the simulation was bound strictly by its fraction of compact to extended sources which determined sources to be variable or non-variable respectively. It was found that the source size distribution, as well as the source count function, were not sufficient to generate realistic source populations, specifically for low-frequency compact sources. At current time there is a dearth of information regarding the distribution of source sizes as a function of flux and frequency to properly generate sources for this simulation, as more data becomes available future work can be to update and further improve this simulation. While difficulties were found for simulating the variability, we were able to determine the required point source population for our simulation to match observed surveys, as

well as limiting the source size of each population based on the IPS limit and the critical scale angular determined by RISS19 for each survey area. We find that: 0.09% of sources must be smaller than 16.3 milli-arcseconds at 185 MHz, 3.51% smaller than 0.053 milli-arcseconds at 1400 MHz, and 19.61% must be smaller than 0.0037 milli-arcseconds at 5500 MHz.

The code used within this thesis is can be found at github.com/PaulHancock/RISS19.

Bibliography

- Antonucci R., 1993, *ARA&A*, **31**, 473
- Armstrong J. W., Rickett B. J., Spangler S. R., 1995, *ApJ*, **443**, 209
- Bannister K. W., et al., 2019, *Science*, **365**, 565
- Becker R. H., White R. L., Helfand D. J., 1995, *AJ*, **450**, 559
- Bell M. E., et al., 2011, *MNRAS*, **415**, 2
- Bell M. E., Huynh M. T., Hancock P., Murphy T., Gaensler B. M., Burlon D., Trott C., Bannister K., 2015, *MNRAS*, **450**, 4221
- Bhat N. D. R., Cordes J. M., Camilo F., Nice D. J., Lorimer D. R., 2004, *ApJ*, **605**, 759
- Bignall H. E., et al., 2003, *AJ*, **585**, 653
- Bright J. S., et al., 2020, *Nature Astronomy*, **4**, 697
- CHIME/FRB Collaboration et al., 2019, *ApJ*, **885**, L24
- Chatterjee S., et al., 2017, *Nature*, **541**, 58
- Chhetri R., Morgan J., Ekers R. D., Macquart J. P., Sadler E. M., Giroletti M., Callingham J. R., Tingay S. J., 2018a, *MNRAS*, **474**, 4937
- Chhetri R., Ekers R. D., Morgan J., Macquart J. P., Franzen T. M. O., 2018b, *MNRAS*, **479**, 2318
- Corbel S., Fender R. P., Tomsick J. A., Tzioumis A. K., Tingay S., 2004, *ApJ*, **617**, 1272

- Cordes J. M., Lazio T. J. W., 2002, arXiv e-prints, [pp astro-ph/0207156](#)
- Cordes J. M., Weisberg J. M., Frail D. A., Spangler S. R., Ryan M., 1991, *Nature*, [354](#), [121](#)
- Dennett-Thorpe J., de Bruyn A. G., 2000, *AJ*, [529](#), [L65](#)
- Dennett-Thorpe J., de Bruyn A. G., 2003, *A&A*, [404](#), [113](#)
- Dennison B., Simonetti J. H., Topasna G. A., 1998, *PASA*, [15](#), [147](#)
- Dewdney P. E., Hall P. J., Schilizzi R. T., Lazio T. J. L. W., 2009, *IEEE Proceedings*, [97](#), [1482](#)
- Dhawan V., Mirabel I. F., Rodríguez L. F., 2000, *ApJ*, [543](#), [373](#)
- Elitzur M., 2012, *ApJ*, [747](#), [L33](#)
- Fender R. P., Belloni T. M., Gallo E., 2004, *MNRAS*, [355](#), [1105](#)
- Ferrière K. M., 2001, *Rev. Mod. Phys.*, [73](#), [1031](#)
- Finkbeiner D. P., 2003, *ApJS*, [146](#), [407](#)
- Frail D. A., et al., 1994, *ApJ*, [437](#), [L43](#)
- Franzen T. M. O., Vernstrom T., Jackson C. A., Hurley-Walker N., Ekers R. D., Heald G., Seymour N., White S. V., 2019, *PASA*, [36](#), [e004](#)
- Gaustad J. E., McCullough P. R., Rosing W., Van Buren D., 2001, *PASP*, [113](#), [1326](#)
- Haffner L. M., Reynolds R. J., Tufte S. L., 1998, *ApJ*, [501](#), [L83](#)
- Haffner L. M., Reynolds R. J., Tufte S. L., Madsen G. J., Jaehnig K. P., Percival J. W., 2003, *ApJS*, [149](#), [405](#)
- Haffner L. M., Reynolds R. J., Madsen G. J., Hill A. S., Barger K. A., Jaehnig K. P., Mierkiewicz E. J., Percival J. W., 2010, in American Astronomical Society Meeting Abstracts #215. p. 415.28

- Hales S. E. G., Riley J. M., WalDRAM E. M., Warner P. J., Baldwin J. E., 2007, [MNRAS](#), **382**, 1639
- Hancock P. J., Drury J. A., Bell M. E., Murphy T., Gaensler B. M., 2016, [MNRAS](#), **461**, 3314
- Hancock P. J., Trott C. M., Hurley-Walker N., 2018, [PASA](#), **35**, e011
- Hancock P. J., Charlton E. G., Macquart J.-P., Hurley-Walker N., 2019, arXiv e-prints, p. [arXiv:1907.08395](#)
- Heeschen D. S., Krichbaum T., Schalinski C. J., Witzel A., 1987, [AJ](#), **94**, 1493
- Hodge J. A., Becker R. H., White R. L., Richards G. T., 2013, [ApJ](#), **769**, 125
- Hopkins A. M., Afonso J., Chan B., Cram L. E., Georgakakis A., Mobasher B., 2003, [AJ](#), **125**, 465
- Hovatta T., Valtaoja E., Tornikoski M., Lähteenmäki A., 2009, [A&A](#), **494**, 527
- Hunstead R. W., 1972, , **12**, 193
- Hurley-Walker N., et al., 2016, [MNRAS](#), **464**, 1146
- Intema H. T., Jagannathan P., Mooley K. P., Frail D. A., 2017, [A&A](#), **598**, A78
- Jansky K. G., 1933, [Nature](#), **132**, 66
- Jauncey D. L., Macquart J. P., 2001, [A&A](#), **370**, L9
- Jauncey D. L., Kedziora-Chudczer L. L., Lovell J. E. J., Nicolson G. D., Perley R. A., Reynolds J. E., Tzioumis A. K., Wieringa M. H., 2000, in Hirabayashi H., Edwards P. G., Murphy D. W., eds, *Astrophysical Phenomena Revealed by Space VLBI*. pp 147–150
- Johnston S., et al., 2007, [PASA](#), **24**, 174
- Jonas J., MeerKAT Team 2016, in *MeerKAT Science: On the Pathway to the SKA*. p. 1
- Kalogera V., Baym G., 1996, [ApJ](#), **470**, L61

- Kedziora-Chudczer L., Jauncey D. L., Wieringa M. H., Walker M. A., Nicolson G. D., Reynolds J. E., Tzioumis A. K., 1997, *ApJ*, **490**, L9
- Kedziora-Chudczer L., Jauncey D. L., Wieringa M. A., Tzioumis A. K., Bignall H. E., 2001, *Ap&SS*, **278**, 113
- Kellermann K. I., Pauliny-Toth I. I. K., 1969, *ApJ*, **155**, L71
- Kellermann K. I., et al., 2007, *Ap&SS*, **311**, 231
- Lien A., Chakraborty N., Fields B. D., Kemball A., 2011, *ApJ*, **740**, 23
- Lorimer D. R., 2008, *Living Reviews in Relativity*, **11**, 8
- Lorimer D., Kramer M., 2005, Handbook of pulsar astronomy. Cambridge observing handbooks for research astronomers, Cambridge Univ. Press, Cambridge, <http://cds.cern.ch/record/1514506>
- Macquart J.-P., Jauncey D. L., 2002, *ApJ*, **572**, 786
- Macquart J.-P., Koay J. Y., 2013, *ApJ*, **776**, 125
- Margalit B., Metzger B. D., 2017, *ApJ*, **850**, L19
- Margalit B., Metzger B. D., Berger E., Nicholl M., Eftekhari T., Margutti R., 2018, *MNRAS*, **481**, 2407
- Mauch T., Murphy T., Buttery H. J., Curran J., Hunstead R. W., Piestrzynski B., Robertson J. G., Sadler E. M., 2003, *MNRAS*, **342**, 1117
- McCallum J. N., Ellingsen S. P., Lovell J. E. J., Phillips C. J., Reynolds J. E., 2009, *MNRAS*, **392**, 1339
- McCulloch P. M., Ellingsen S. P., Jauncey D. L., Carter S. J. B., Cimò G., Lovell J. E. J., Dodson R. G., 2005, *AJ*, **129**, 2034
- McGilchrist M. M., Baldwin J. E., Riley J. M., Titterton D. J., Waldram E. M., Warner P. J., 1990, *MNRAS*, **246**, 110
- Mooley K. P., Frail D. A., Ofek E. O., Miller N. A., Kulkarni S. R., Horesh A., 2013, *ApJ*, **768**, 165

- Morgan J. S., et al., 2017, *MNRAS*, 473, 2965
- Narayan R., Blandford R. D., Hewish A., Lyne A. G., Mestel L., 1992, *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences*, 341, 151
- Netzer H., 2015, *ARA&A*, 53, 365
- Özel F., Freire P., 2016, *ARA&A*, 54, 401
- Pietka M., Fender R. P., Keane E. F., 2014, *MNRAS*, 446, 3687
- Qian S. J., Quirrenbach A., Witzel A., Krichbaum T. P., Hummel C. A., Zensus J. A., 1991, *A&A*, 241, 15
- Rawlings J. I., et al., 2015, *MNRAS*, 452, 4111
- Rickett B. J., 1986, *ApJ*, 307, 564
- Rickett B. J., 1990, *ARA&A*, 28, 561
- Rickett B., 2002, *PASA*, 19, 100
- Rickett B. J., Quirrenbach A., Wegner R., Krichbaum T. P., Witzel A., 1995, *A&A*, 293, 479
- Rickett B. J., Witzel A., Kraus A., Krichbaum T. P., Qian S. J., 2001, *ApJ*, 550, L11
- Rickett B. J., Kedziora-Chudczer L., Jauncey D. L., 2002, *ApJ*, 581, 103–126
- Ruiz M., Shapiro S. L., Tsokaros A., 2018, *Phys. Rev. D*, 97, 021501
- Sabater J., et al., 2019, *A&A*, 622, A17
- Saripalli L., Subrahmanyam R., Thorat K., Ekers R. D., Hunstead R. W., Johnston H. M., Sadler E. M., 2012, *ApJS*, 199, 27
- Scheuer P. A. G., 1968, *Nature*, 218, 920
- Seymour N., et al., 2020, *PASA*, 37, e013

- Shimwell T. W., et al., 2019, *A&A*, **622**, [A1](#)
- Subrahmanyan R., Ekers R. D., Saripalli L., Sadler E. M., 2010, *MNRAS*, **402**, [2792](#)
- Taylor J. H., Cordes J. M., 1993, *ApJ*, **411**, 674
- Thyagarajan N., Helfand D. J., White R. L., Becker R. H., 2011, *ApJ*, **742**, 49
- Tingay S. J., Hancock P. J., 2019, *AJ*, **158**, [31](#)
- Tingay S. J., et al., 2013, *PASA*, **30**, [e007](#)
- Valtaoja E., Savolainen T., Wiik K., Lähteenmäki A., 2003, in Zensus J. A., Cohen M. H., Ros E., eds, *Astronomical Society of the Pacific Conference Series Vol. 300, Radio Astronomy at the Fringe*. p. 169
- Walker M. A., 1998, *MNRAS*, **294**, 307
- Wenger M., et al., 2000, , [143](#), [9](#)
- Williams W. L., et al., 2016, *MNRAS*, **460**, [2385](#)
- Williams W. L., et al., 2018, *MNRAS*, **475**, [3429](#)
- Windhorst R., Mathis D., Neuschaefer L., 1990, in Kron R. G., ed., *Astronomical Society of the Pacific Conference Series Vol. 10, Evolution of the Universe of Galaxies*. pp 389–403
- Yao J. M., Manchester R. N., Wang N., 2017, *ApJ*, **835**, 29
- van Haarlem M. P., et al., 2013, *A&A*, **556**, [A2](#)

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.