

Abstract

To autonomously detect Type III Solar Radio Burst (SRBs) observations taken with the Irish Low Frequency Array (ILofar) are binary masked with an optimal threshold method, determined to be the mean method, and ran through a shape detecting transform to locate the SRBs in the radio spectra. Once located the detected SRBs initial and final frequencies are then fed into a coronal density model to determine their electron density change across this bandwidth and thus their changing distance from the centre of The Sun. This bandwidth and distance change allows for the determination of the SRB drift rate, calculated to be 9 MHz/s, and the SRB velocity, calculated to be 0.3c.

1 Introduction

The Sun is made up of a charged particle soup, separated electrons and ions, called plasma which is produced by nuclear fusion at its center. The movement of these charged particles induces a magnetic field, which has a feedback effect changing how the particles move inducing another magnetic field. The repetition of this process produces the suns entire magnetic field and is called the solar dynamo.

Solar areas with strong magnetic fields are called active regions and they are where phenomena such as solar flares, coronal mass ejections (CMEs), and solar radio bursts (SRBs) are produced. These active regions can be identified by dark patches on the solar surface, photosphere, called sunspots and are only dark due to the contrast between it and the brighter surrounding area. Sunspots can last days, weeks, or months before disappearing and can be seen in the Magnetogram images shown on the left in Figure 7 where the black regions are inward directed magnetic fields and the whiter regions are outward directed magnetic fields.

When electrons from solar active regions are ejected along these magnetic field lines they excite the solar plasma through bump on the tail instability at the plasma frequency, F_p , or its harmonic, $2F_p$. As these electrons travel through regions of decreasing density, and thus increasing their height from the centre of the sun, they produce lowering frequencies which can be observed as SRBs by the Low Frequency Array (LOFAR) as seen in the Radio Spectra shown on the left in Figure 9. These SRBs are quite uniform in their shape on radio spectra and have the potential to be detected autonomously with a shape detecting algorithm which is what this investigation will attempt to do.

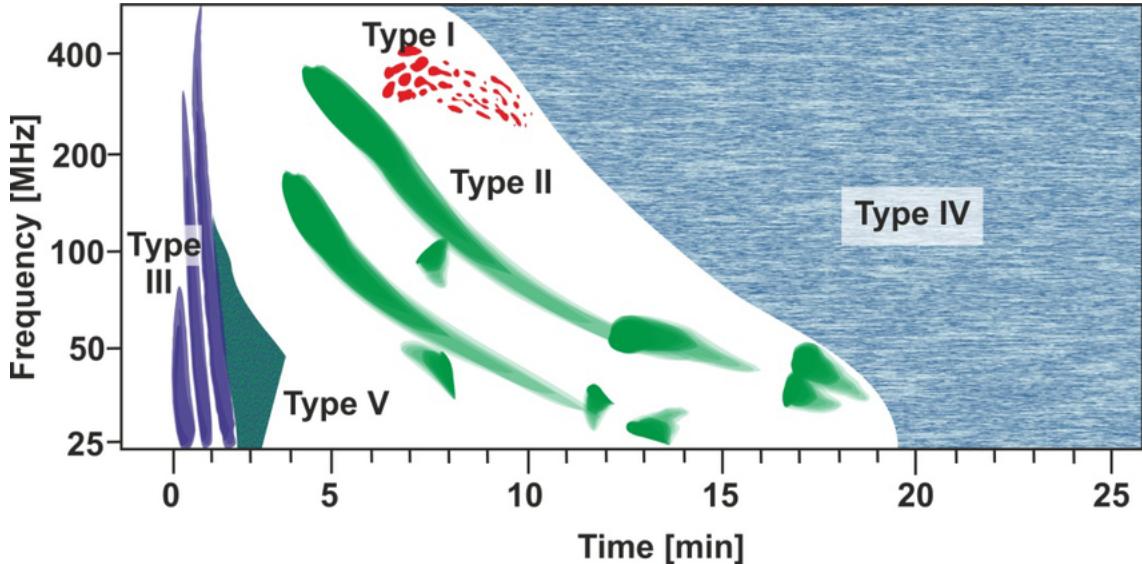


Figure 1: Classification of Solar Radio Bursts

There are five types of SRBs, Type I through to Type V as shown above in Figure 1. Type I are one second bursts that usually occur in large numbers in the 80 - 200 MHz band. Types II are slow frequency drift bursts that happen over longer periods of time, typically between 3-30 minutes over 20 - 150 MHz band. Type III are fast frequency drift bursts that usually last for a few seconds but groups can last for minutes and are typically picked up at 10 kHz - 1 GHz range. Type IV have three distinct groups; stationary, moving, and flare, and are characterised by their continuums and found between 20 MHz - 2 GHz. And finally Type V which are short lived continuums that follow type III bursts and last for a few minutes. This investigation will focus only on the automated detection of Type III radio bursts.

2 Background

A technique developed by Lobzin et al. (2009) discusses how to write an algorithm that can detect radio bursts automatically. Lobzin does this by first cleaning the dynamical spectrum to remove radio frequency interference (RFI) and applying a smoothing filter to improve the signal to noise ratio. Lobzin then creates a binary mask by using a threshold method to set pixel values to 1 for all the local maxima along time axis at each frequency and all other pixels set to a value of 0. This forms a straight line where the Type III radio burst is. Finally he uses a shape detecting Radon transform to detect the straight lines which get recorded. To test Lobzin's algorithm it was reproduced with python with the same threshold method and transform. To optimise the algorithm multiple threshold methods were tested along with different shape detecting transforms. Furthermore the addition of a coronal density model meant that the height change and velocity of the radio bursts could now also be determined.

2.1 Data Thresholding

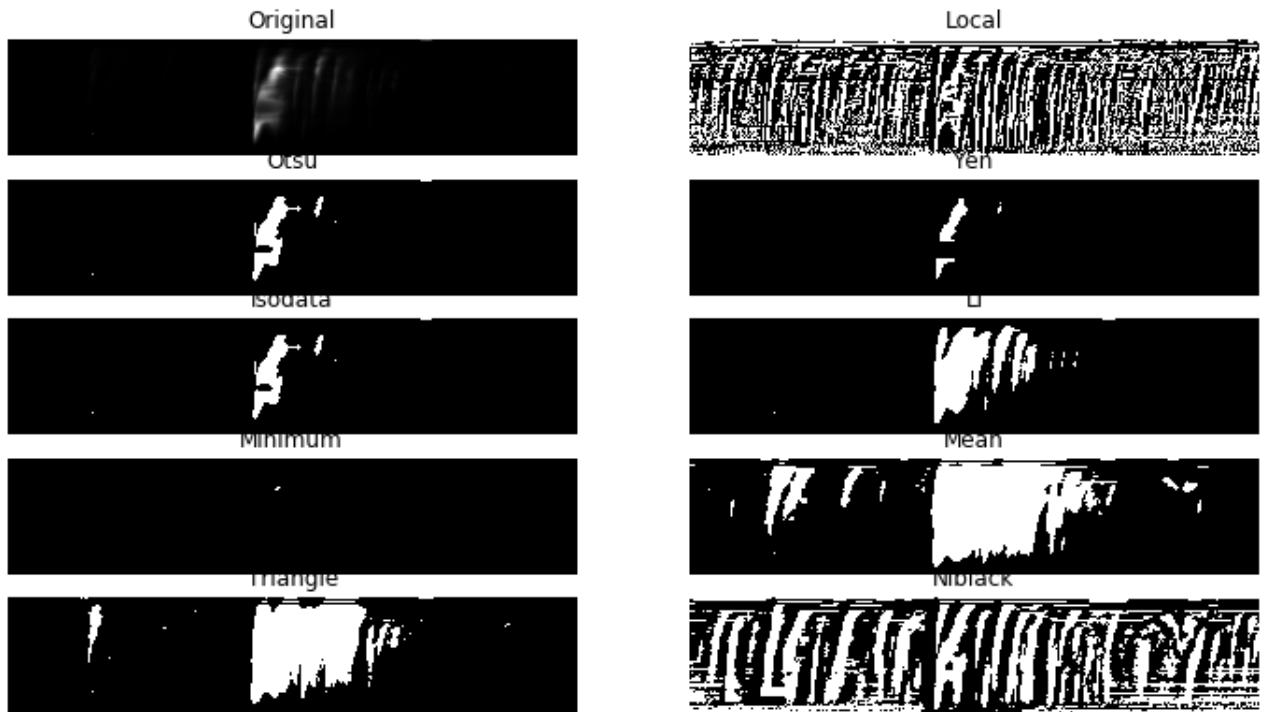


Figure 2: Comparison of threshold methods on 2nd September 2017 radio spectrum

The threshold method used by Lobzin in his process is the Otsu threshold method which determines a threshold value by plotting the foreground and background pixels in a histogram which shows the optimal threshold separation of the two pixel classes. However there are other threshold methods that can be used as seen in Figure 2. The Local threshold value is the weighted mean of intensities for the local neighbourhood of pixels. The Yen Method determines a value by treating the foreground and background pixels as two different signal sources that are optimally thresholded when the sum of the two class entropies reaches its maximum.

Isodata stands for Iterative Self Organizing Data, and as such it iterates over the data, dividing it up into pixel clusters, re-merging it, and dividing it up into different clusters over and over until it reaches a point at which the maximum number of iterations is reached and a threshold is determined by the average distance between cluster centers. The Li Method determines a threshold by determining the minimum cross entropy between pixel classes. The Minimum Method plots a histogram of the data and smooths it until there are only two maxima, with the minima being the threshold value. The Mean threshold value is the mean of grayscale plot.

The Triangle Method encloses a histogram of the pixels with a right angled triangle such that the hypotenuse touches the top of the highest bin and the top of the furthest bin. The threshold is then determined by the bin the has the largest distance from the hypotenuse along a perpendicular line to the hypotenuse plus an offset value. Finally the Niblack method determines threshold values for every pixel by using specific formulae that uses the mean and standard deviation of the local neighbourhood defined by a square window centered around the pixel.

Plotting a cross section of the radio spectrum observed on 2nd September 2017 showing the intensity values at a given frequency (55 Mhz) as a function of time can demonstrate the detail loss at different threshold values. As seen in Figure 3 certain threshold values cut off too little, like the Triangle, Li, and Niblack Methods, thus the data isn't clean enough to use the transforms effectively. Whereas other methods like the Minimum, Yen, Otsu, and Isodata cut off so much that required finer detail is lost.

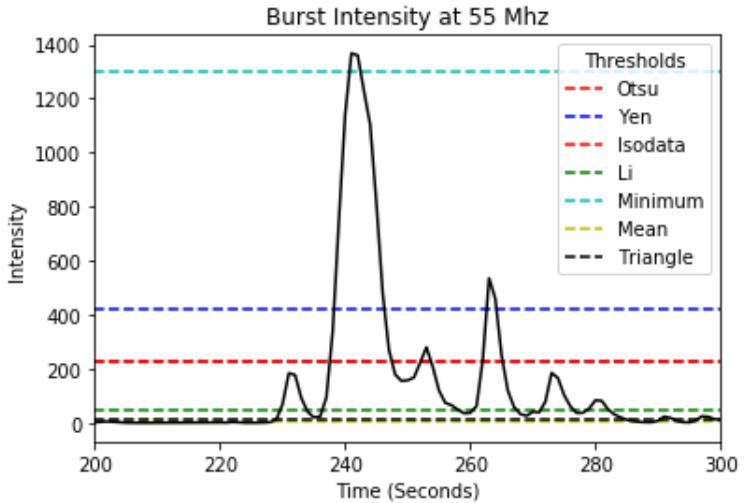


Figure 3: Burst intensity at 55 Mhz on 2nd September 2017 radio spectrum with multiple threshold method values showing where each method cuts the raw data when thresholding

Comparing where the methods cut the data in Figure 3 and the binary output of the data shown in Figure 2 demonstrates how different methods produce different spectra. The Minimum method cuts so high that the spectrum only has a single white dot. Whereas the Niblack method cuts so low that the spectrum is streaked with white lines that do not match the SRB in the original spectrum due to the boosting of the values of negligible radio signals. However the Mean threshold method cuts high enough to ignore the negligible radio signals, but low enough to keep other necessary SRB signals, as such this method is the most optimal.

2.2 Shape Detection

For shape detection Lobzin uses the Radon Transform, which in a two-dimensional case is able to locate shapes by determining parameters r and θ , where r is the distance of the line from the origin of the (x, y) plane and θ is the angle of the line from the vertical axis. An image represented by a function $\mathcal{F}(x, y)$ can be defined as a series of line integrals through $\mathcal{F}(x, y)$ projected at different offsets from the origin via the Radon Transform described by:

$$\mathcal{R}(r, \theta) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathcal{F}(x, y) \mathcal{G}(r, \theta) dx dy \quad (1)$$

Where $\mathcal{G}(r, \theta)$ is $\delta(x \cos \theta + y \sin \theta - r)$. In the formula above the function $\mathcal{F}(x, y)$ is convolved with the delta function $\mathcal{G}(r, \theta)$ rotating about a series of angles. This produces a collection of values $\mathcal{R}(r, \theta)$ at all θ per point (x, y) . Plotting \mathcal{R} against θ produces a transform image in which the sinusoids produced by rotating around each pixel will all intersect at a specific pixel distance and angle which is where the line is located in the image $\mathcal{F}(x, y)$.

The Hough Transform produces the same result but is more straightforward in its application. A straight line, $y = mx + c$, represented by an intercept, c , and slope, m , in parameter space gives rise to unbounded values of the slope. So instead it's better to describe this line in polar coordinates, $r = x \cos \theta + y \sin \theta$. As shown in Figure 4 taking the image of a dot as a substitute for a single pixel we can see how the Hough transform is able to identify the shape by rotating around all possible lines (in red) that can go through the dot producing a transform image of a single sinusoid by plotting the angles θ against the distance r . Now taking an image of a single line made up of multiple pixels the transform rotates through all possible lines for the pixels, producing a transform image with a series of sinusoids each representing a pixel. All sinusoids intersect at two points, 90° and -90° , exactly 180° apart thus determining one line. Finally taking two lines, the transform image shows an intersection again at $\pm 90^\circ$ and another at approximately -43° locating the existence of two lines at two distinct distances and angles.

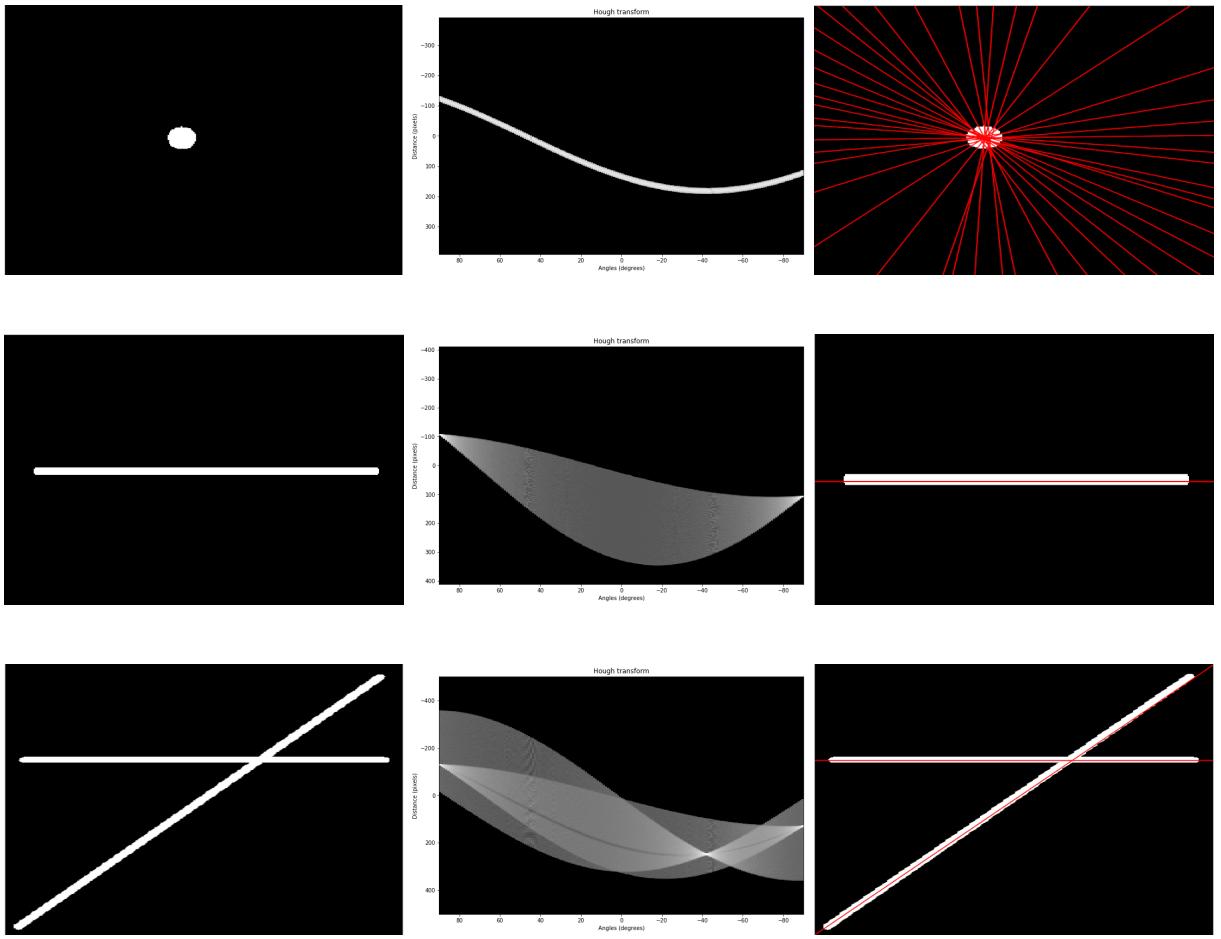


Figure 4: Right: Binary images of a dot, a line, and two intersecting lines. Middle: The Hough Transform image associated with each shape showing a series of sinusoids representing each pixel intersecting at a specific angle and distance which allows for the identification of the shape. Left: Post transform image showing original image with detected lines.

2.3 Newkirk's Coronal Density Model

Coronal density models can be used to infer the drift velocity of radio burst once the frequency drift is known. Firstly the start frequency and the end frequency of the drift is used to determine electron density using Equation 3, then using the Newkirk density model, shown in Figure 5 which shows the trend in the Newkirk equation for two regions, a conversion from the electron density to height can be made. Given that the time between the two frequencies is known, and now the distance travelled (difference in heights) is known, a value for the drift velocity can be determined.

Newkirk Equation:

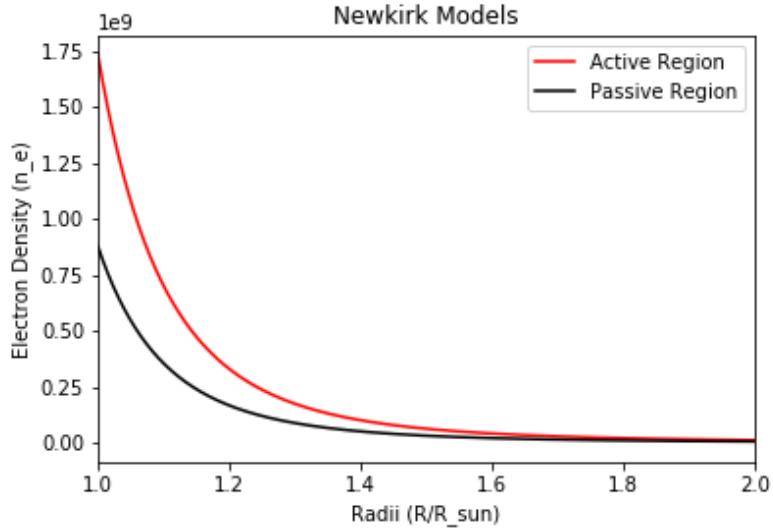


Figure 5: Passive and Active Newkirk Models

$$n_e(R) = f_0 n_0 10^{4.32 \frac{R_{sun}}{R}}$$

$$\therefore R = \frac{4.32 R_{sun}}{\log \frac{n_e}{n_0}} \quad (2)$$

Where f_0 is the fold value with $f_0=1$ for passive regions and $f_0=2$ for active regions , n_e is the electron density, n_0 is $4.2 \cdot 10^4$, and R_{sun} is $6.957 \cdot 10^{10}$ cm.

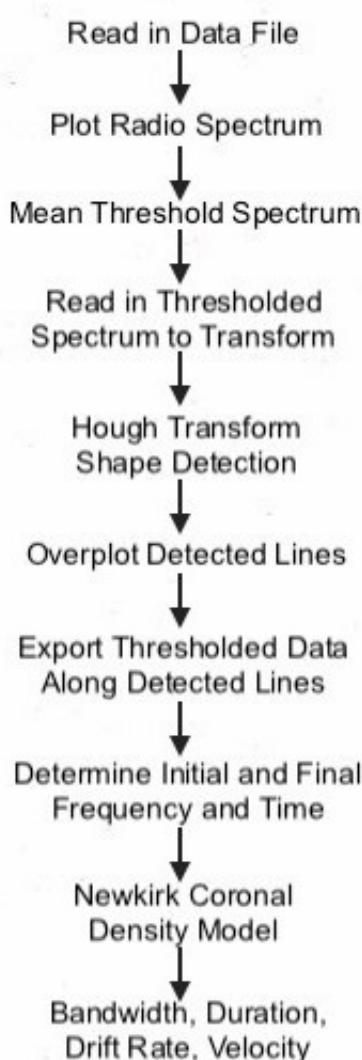
$$\omega = \sqrt{\frac{n_e e^2}{m_e \epsilon_0}} = 2\pi f$$

$$\therefore n_e = \frac{4\pi^2 m_e \epsilon_0}{e^2} \cdot f \quad (3)$$

Where m_e , ϵ_0 , and e^2 are known constants. Thus substituting (2) into (1) would produce an equation that gives the Radius (Height) if the frequency is known.

3 Computational Method

As seen throughout Section 2 Lobzin's algorithm can be updated and optimised to better locate SRBs by swapping out the threshold method and shape detection transform used as well as adding a coronal density model to determine SRB height change and velocities. The initially used Otsu threshold method was changed due to it cutting the data so high that it removed too much finer detail from the spectra. Thus the Mean threshold method was selected as it cuts the data high enough to ignore negligible signals and low enough to keep required SRB signals as shown in Section 2.1. The Radon transform was also swapped out for the Hough transform due to the latter being more straightforward in application when determining straight lines and easier at working backwards when determining initial and final frequency and time values.



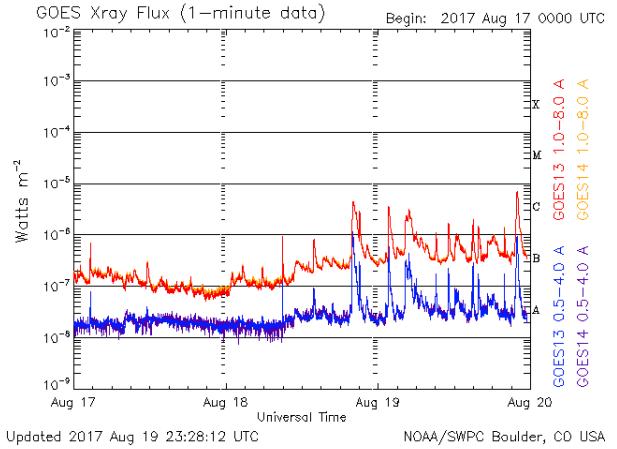
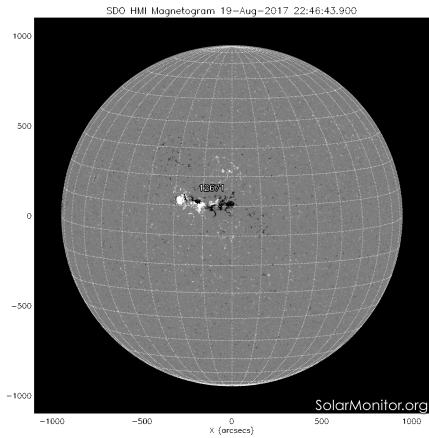
As seen in Figure 6 the code, written in Python, first reads in the Lofar data file. Once the spectrum is plotted the Mean threshold method is performed which turns the spectrum into a binary image as seen in Figure 9. The binary spectrum is then ran through the Hough transform and the detected lines are overplotted, also shown in Figure 9. Using a ForLoop and the Index function the exported data along the detected lines can be used to determine the initial and final frequency and time. The two frequency values are now fed into the Newkirk coronal density model which determines the change in electron density and thus the electron height change in the sun. The height change and the time difference can now be used to determine the electron velocity, these values are displayed in Table 1 and the distribution of these values are plotted in Figure 12.

Figure 6: Algorithm Flowchart

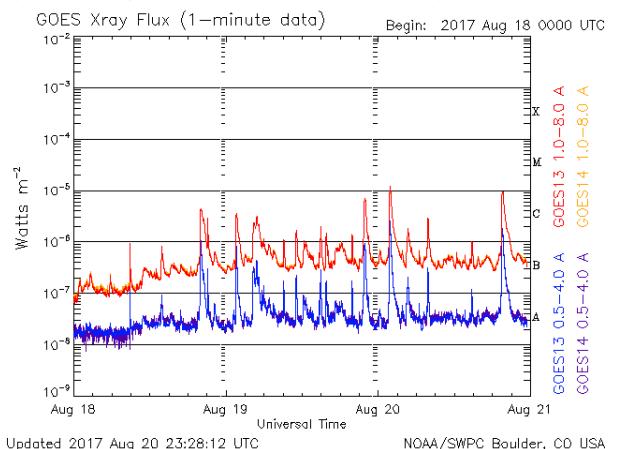
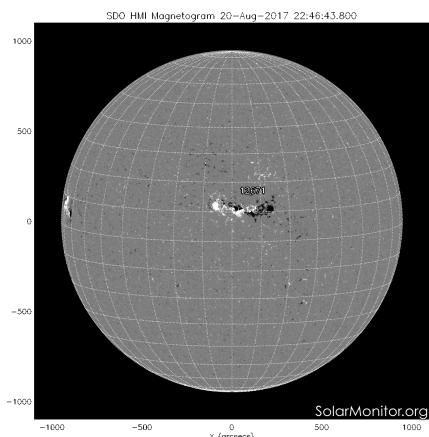
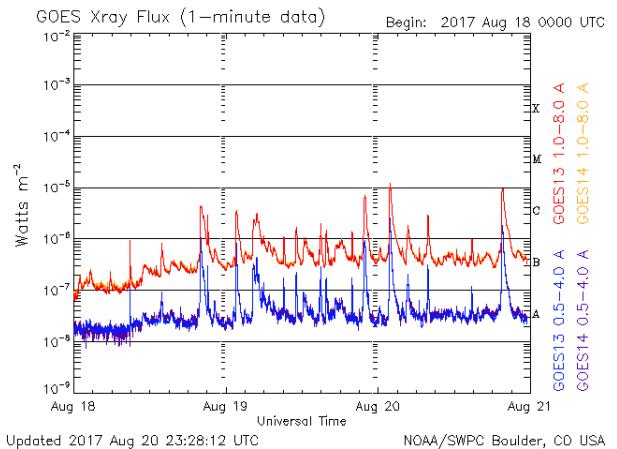
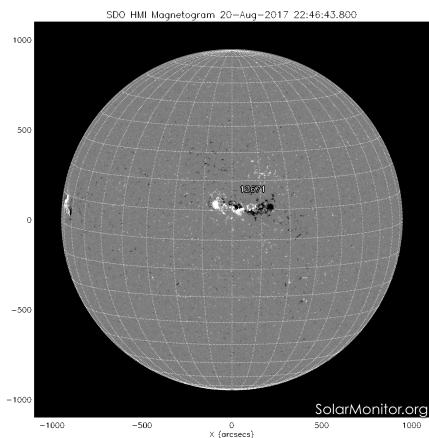
4 Results and Discussion

4.1 Magnetograms and X-Ray Flux Data

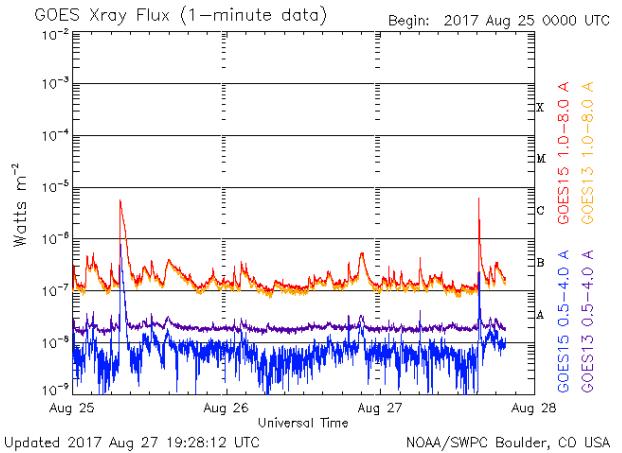
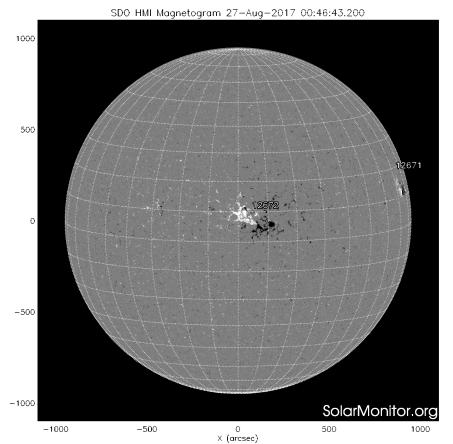
19th August 2017



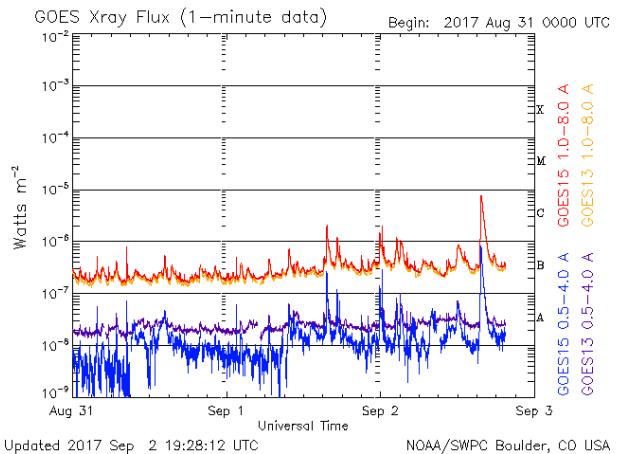
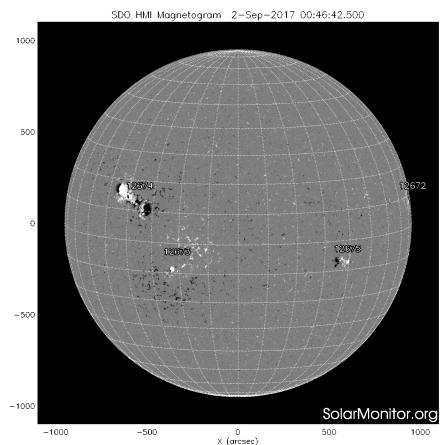
20th August 2017



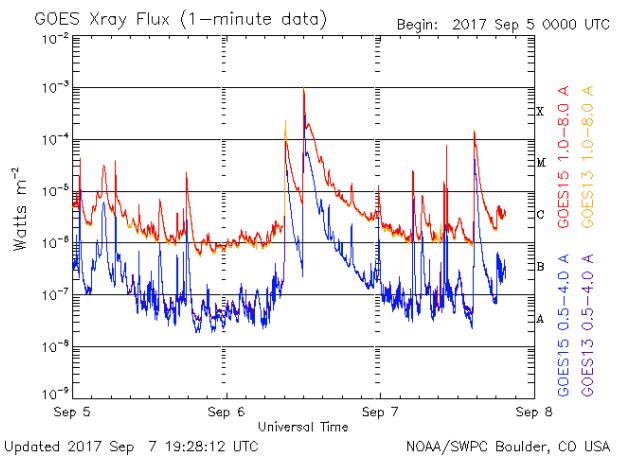
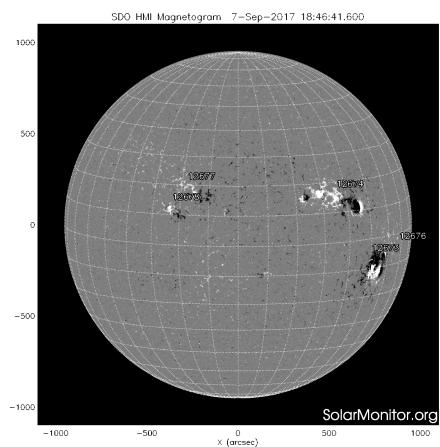
27th August 2017



2nd September 2017



7th September 2017



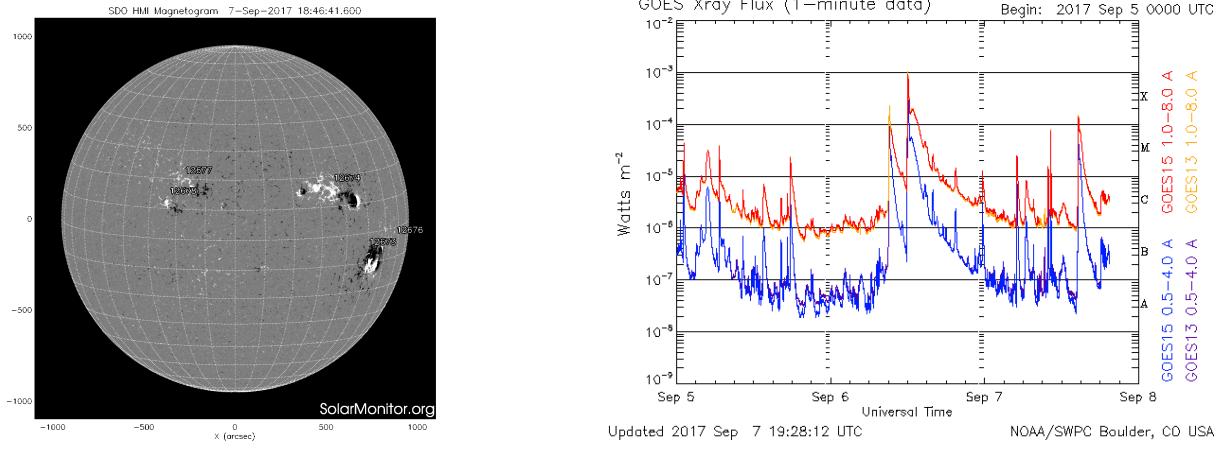


Figure 7: Left: Active magnetic regions observed with the Solar Dynamics Observatory (SDO) Helioseismic and Magnetic Imager (HMI). Right: Solar X-Ray Flux observed with the Geostationary Operational Environmental Satellite (GOES) showing solar flares.

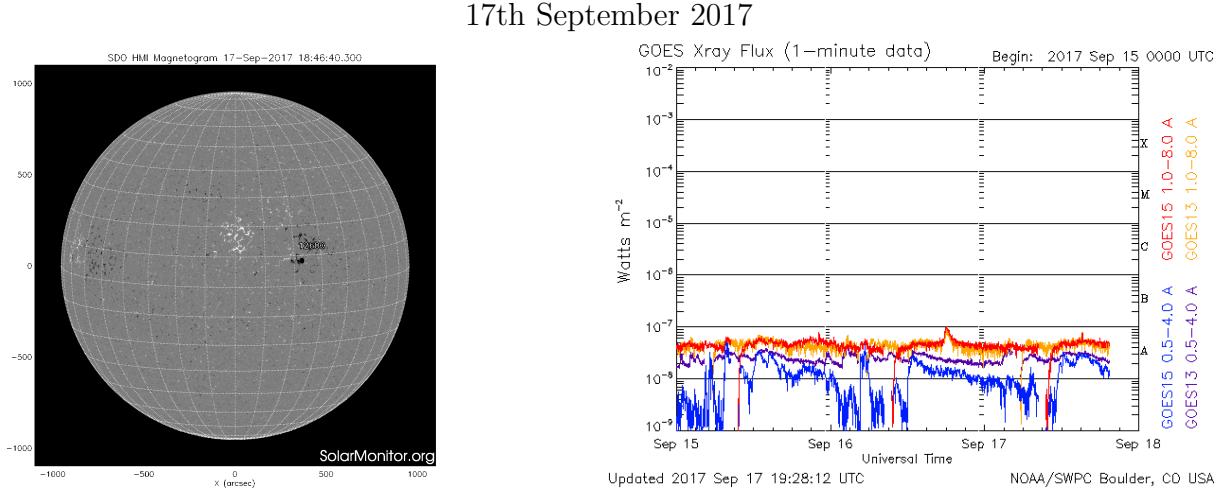


Figure 8: Magnetogram and X-Ray Flux showing little to no sunspots or flares resulting in no viable burst data at radio

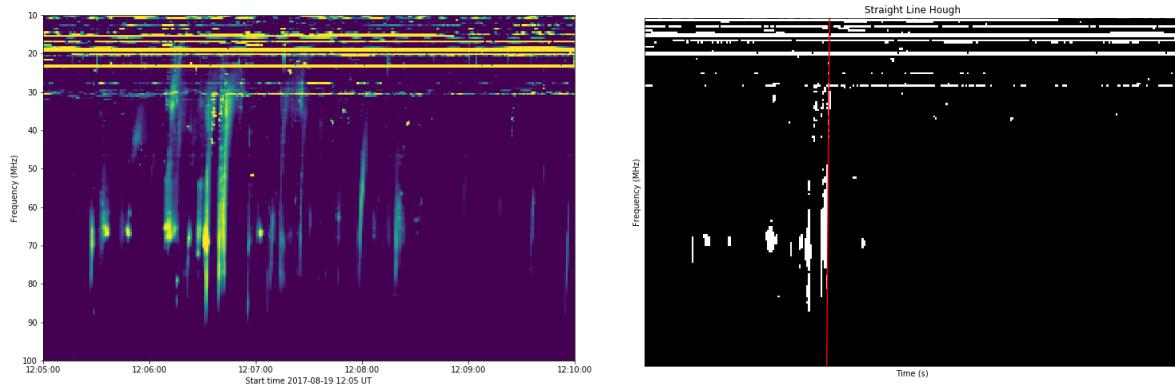
Since SRBs are produced in solar active regions and alongside solar flares it's prudent to demonstrate that during the times the radio spectra shown in Figure 9 are taken that there is active regions on the solar surface. On the left in Figure 7 the SDO HMI Magnetograms show sunspots on the photosphere when all radio spectra are taken where the black regions are inward directed magnetic fields and the white regions are outward directed magnetic fields and are indicative of active regions.

On the right side of Figure 7 GOES X-Ray Flux data, taken during the same time as the spectra in Figure 9, track solar activity and monitor for solar flares which are also indicative of active regions and can be associated with SRBs. The combination of both the Magnetograms and X-Ray Flux data is thus able to demonstrate that during the times the SRBs are observed active regions are present on the solar surface.

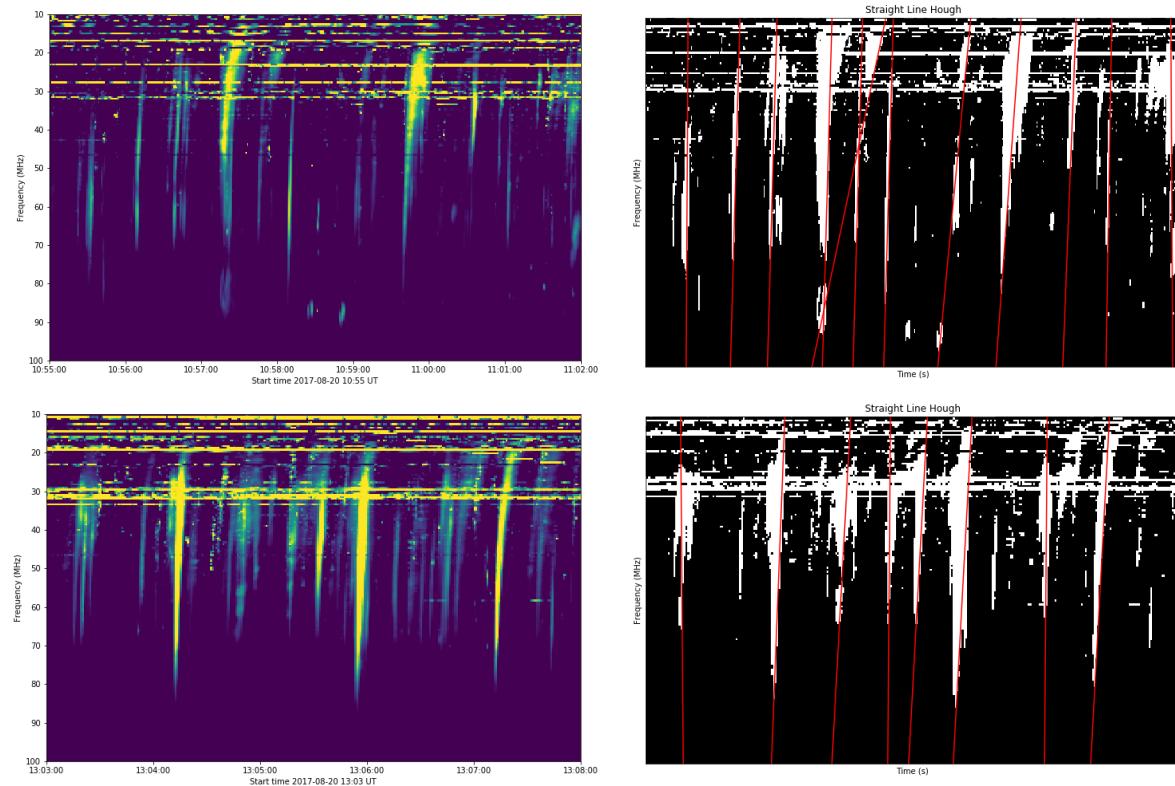
While Lofar radio data for the 17th September is taken when ran through the algorithm discussed in Section 3 it produced no bursts. When the Magnetogram and X-Ray Flux shown in Figure 8 are later investigated it can be seen that the magnetic fields aren't as large as in previous images and the flux is lower than in previous graphs as such there is no observed burst found in the radio data.

4.2 Thresholding and Burst Detection

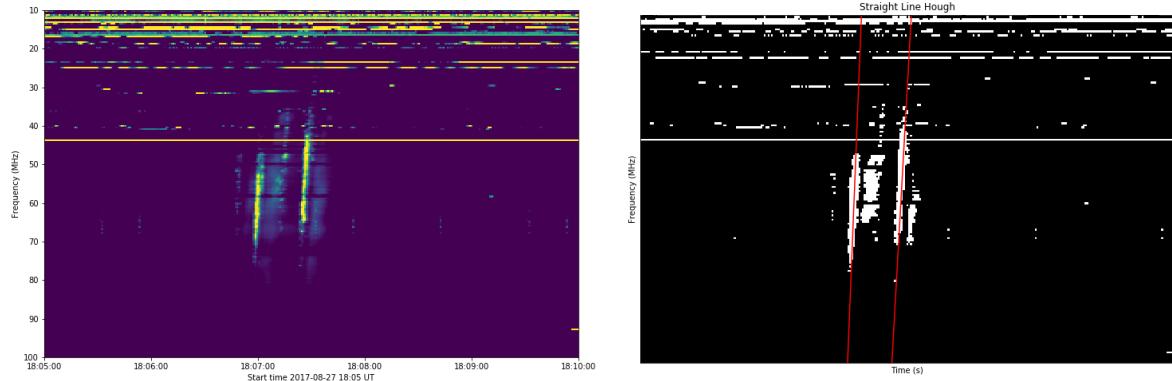
19th August 2017



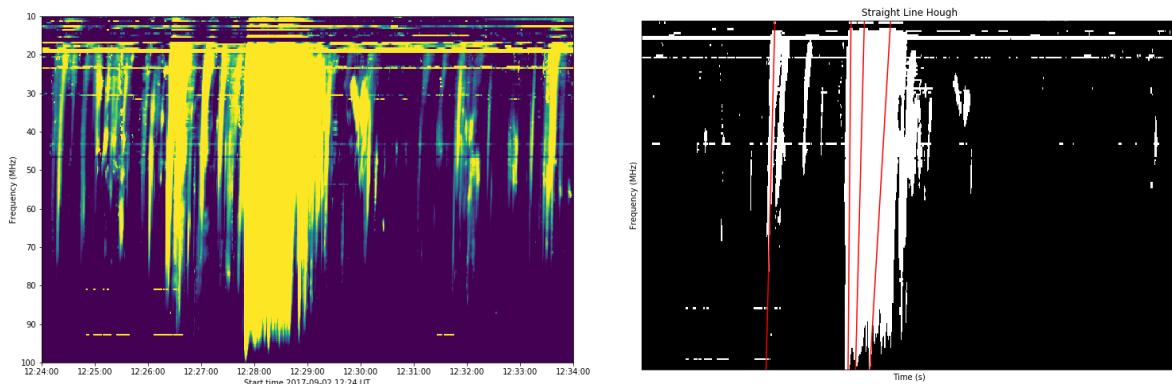
20th August 2017



27th August 2017



2nd September 2017



7th September 2017

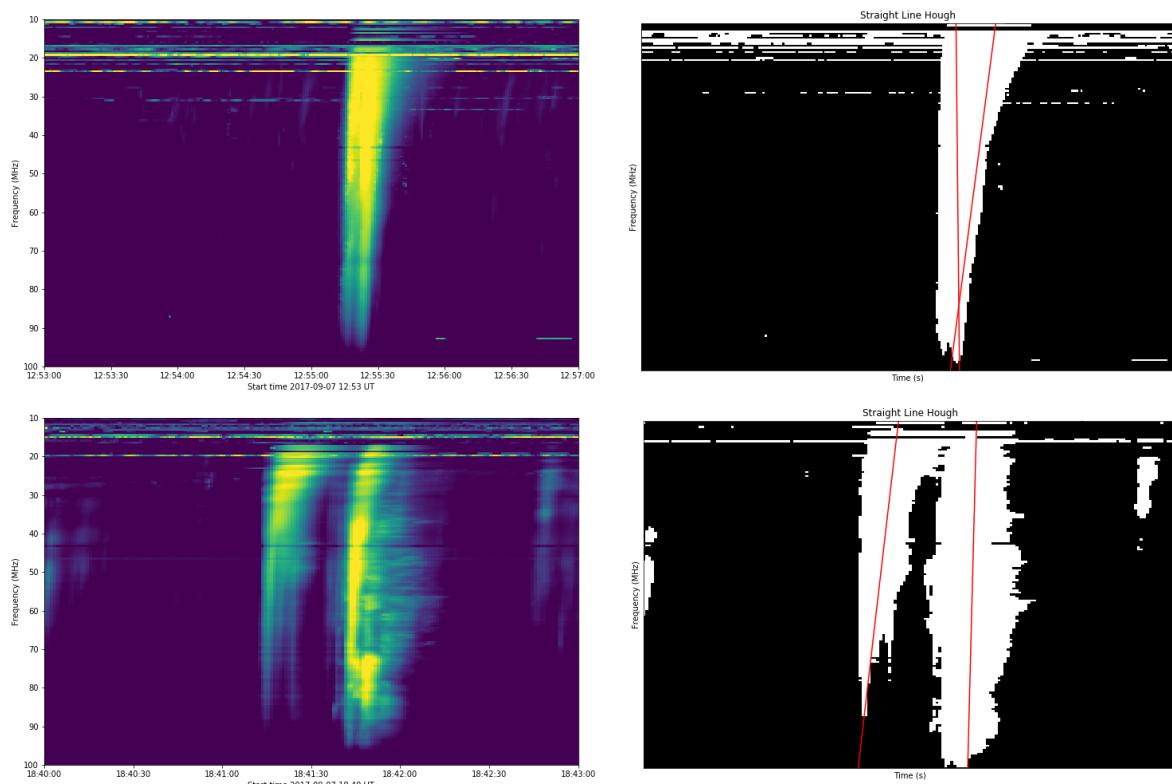


Figure 9: Left: Radio Spectra from the Irish Lofar station showing solar radio bursts. Right: Spectra post mean thresholding and Hough Transform shape detection with red lines indicating the bursts detected.

As seen above in Figure 9 radio spectra observed with Lofar have been plotted, mean thresholded and straight line detected with the Hough transform using the method outlined in Section 3. As shown in the above Section, Section 4.2, all Type III's were observed during an active region being present on the solar surface. On the left in Figure 9 the detected Type III's are indicated with an overplotted red line with the properties of these burst shown in Table 1 below.

The effectiveness of the line detection is dependent on the discreteness of the bursts. This can be seen above where the Type III bursts present in the spectra for the 19th, 20th, and 27th August are well defined and have little Type V continuums thus can be straight line detected independently of each other. However the burst seen in the spectra from the 2nd and 7th September aren't as easy to detect properly due the diffuse nature of the bursts such that multiple bursts bleed into one another due to their overlapping continuums because the threshold value is too low, as evident by the yellow line in Figure 10, thus the two bursts can not be differentiated easily. This causes the detected red lines to veer into the middle of the burst mistaking the Type V continuum as being part of the Type III burst and thus recording incorrect burst durations because the lines are no longer reading just the Type III's.

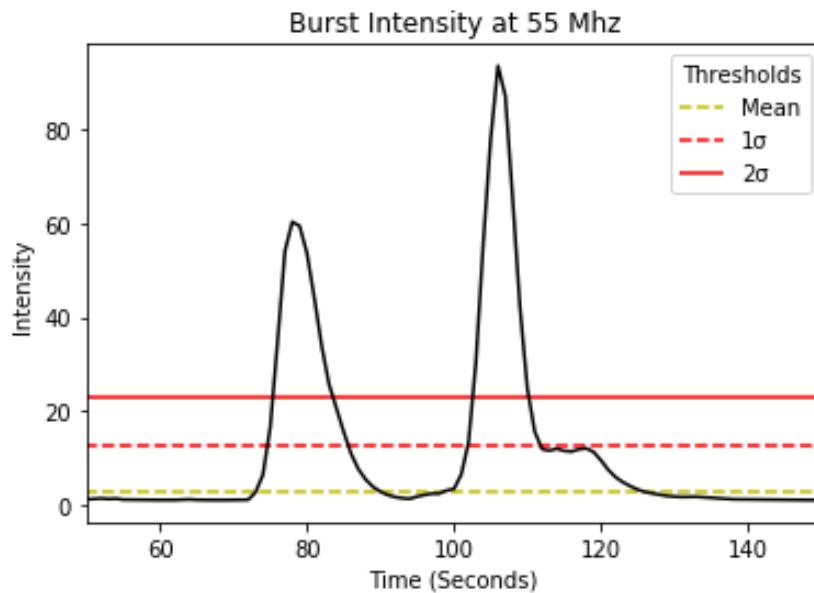


Figure 10: Burst intensity at 55 MHz on 7th September 2017. The plot shows Intensity (Y-axis, 0 to 80) versus Time (Seconds) (X-axis, 60 to 140). A black line represents the burst intensity. Three horizontal lines represent thresholds: Mean (yellow dashed), 1σ (red dashed), and 2σ (red solid). The 2σ threshold captures both peaks of the burst, while the 1σ threshold only captures the second peak.

This can be remedied by thresholding at a slightly higher rate, but this would involve either manually choosing a threshold value for the data or using the standard deviation of the mean threshold value shown in Figure 10 which would produce a new Hough image for the 7th September shown below in Figure 11

7th September 2017

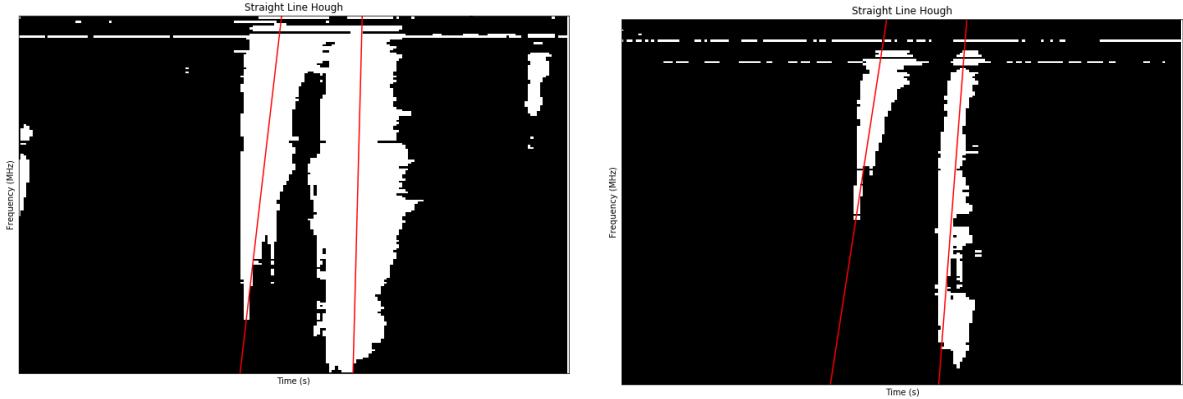


Figure 11: Left: Original Hough transform image for the 7th September. Right: New Hough transform image using standard deviation value 2σ

It's now evident to see that when dealing with spectra that have discrete bursts the mean threshold is effective, however it's effectiveness lessens when dealing with bursts that have larger continuums as the threshold isn't high enough. The solution to this is to use the standard deviation of the mean threshold, however how to automatically determine when to use just the mean and when to use the mean plus it's deviation, thus differentiating between discrete and diffuse images, without manually changing the algorithm still requires work.

A further problem that now arises after fixing the incorrect values of the burst duration is the loss of the true bandwidth when using the mean plus the standard deviation which is evident when comparing the old Hough image with the new image shown in Figure 11. At present some results will always be incorrect in some way since just using the mean threshold value causes the duration to increase, but using the mean plus standard deviation value causes the bandwidth to decrease.

4.3 Burst Properties

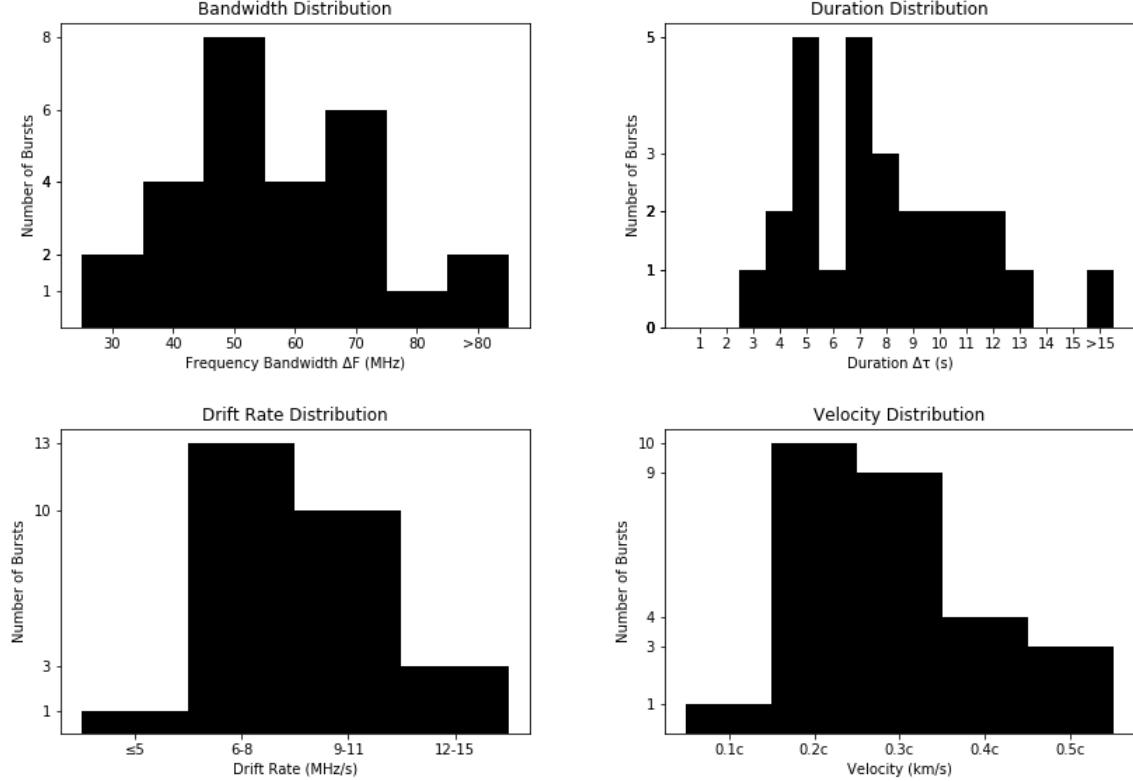


Figure 12: Distribution of the observed radio burst properties

The graphs shown in Figure 12 are the distributions of the properties shown in Table 1 of the bursts identified in Figure 9. These distributions show that the majority of the frequency bandwidths are hovering around 50MHz with a mean value of 63MHz which is expected since Type III's drift rapidly in frequency versus time domain thus have a wide bandwidth. The duration of the radio bursts detected was determined to be approximately 5-7s which is high when compared with White et al. (2007) commonly accepted value of approximately 3s. Reid and Ratcliffe (2014) state that the duration of individual type III bursts varies as function of frequency, however it's more likely that the lines detected in Figure 9 should be at a sharper angles in some cases which would decrease the duration time, this is due to the bleeding of Type V continuums and the Type III bursts. The drift rate distribution mainly appears to be about 7-10 MHz/s with the mean value being 9 MHz/s which is similar to White's value of 10 MHz/s.

Finally using the Active Newkirk model values for the height change are determined as shown in Figure 13 and velocity values are then calculated. The burst velocity distribution then shows that most of the bursts have speeds of $0.2\text{-}0.3c$, where c is the speed of

light, with a mean value

of $0.3c$ which is a little

less than White's value of $0.5c$. However it's also stated in (Benz 2002) that type III radio bursts have velocities ranging from 0.1 to $0.6c$ thus all velocities in Table 1 fall in range of the expected value.

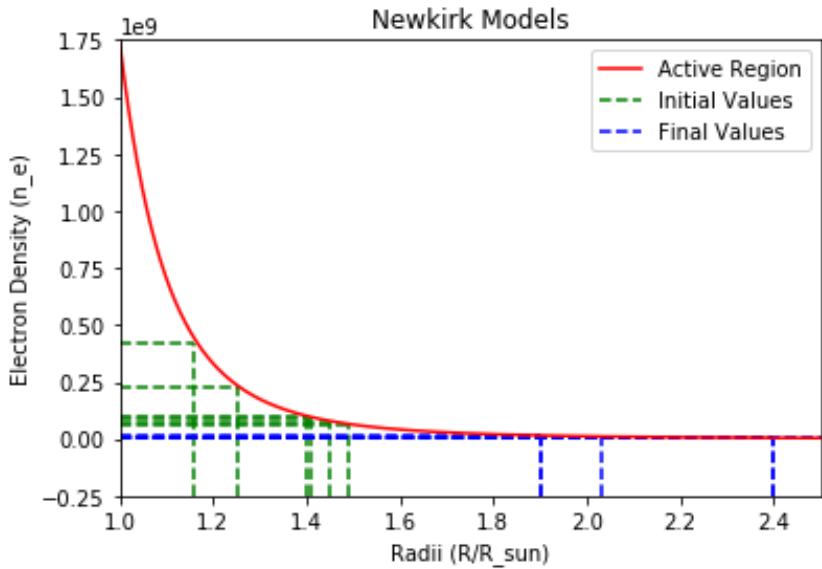


Figure 13: Active Newkirk Model with Initial and Final Values

Table 1: Table of properties from identified bursts shown in Figure 9

Date Time DD/MM/YY	Bandwidth ΔF MHz	Duration $\Delta\tau$ s	Drift Rate MHz s^{-1}	Height Change ΔH R/R _s	Velocity c km s^{-1}
19/08/2017	60	7	9	0.63	0.2
20/08/2017	45	3	15	0.46	0.4
	49	8	6	0.78	0.2
	52	9	6	0.94	0.2
	73	11	7	1.24	0.3
	53	7	8	1.14	0.4
	53	7	8	0.94	0.3
	67	10	7	1.02	0.2
	53	5	11	1.14	0.5
	38	4	10	0.42	0.2
	60	4	15	0.98	0.6
	45	5	9	0.69	0.3
	70	9	8	1.23	0.3
	50	8	6	0.93	0.3
	50	5	10	0.93	0.4
	50	8	6	0.93	0.3
	71	11	6	1.23	0.3
	52	5	10	0.94	0.4
	68	12	6	1.22	0.2
27/08/2017	30	7	4	0.24	0.1
	40	6	7	0.43	0.2
02/09/2017	117	13	9	1.18	0.2
	70	12	6	1.03	0.2
	165	19	9	1.27	0.2
07/09/2017	75	5	15	1.05	0.5
	75	7	11	1.05	0.3
	80	10	8	1.27	0.3

5 Conclusion

Beginning with Lobzin's discussion on the autonomous detection of Type III radio burst the two components of his algorithm, Otsu thresholding and Radon transform shape detection, were tested and compared with other methods and transforms to determine the most optimal technique. As discussed in Section 2.1 the best thresholding was the Mean method as it cuts the data high enough to ignore negligible signals and low enough to keep required SRB signals unlike the Otsu method which cut the data so high that it removed too much of the finer detail. The Radon transform outlined in Section 2.2 was also swapped out as the Hough transform was found to be more straightforward in application when determining straight lines and working backwards to determine frequency and time values. The addition of the Newkirk model, Section 2.3, meant that the height change and velocity of the radio bursts could also now be determined.

The Magnetograms and X-Ray Flux data shown in Section 4.1 demonstrates that during the times the SRBs are observed active regions are present on the solar surface by showing the presence of sunspots and solar flares respectively, both of which are indicative of these regions. The effectiveness of the line detection was shown to be dependent on the threshold value and how independent the bursts are. If a series of bursts are shown to be discrete then the mean threshold on its own is sufficient in aiding in its detection however if the threshold is too low and the bursts are too close together their overlapping continuums can make it more difficult to differentiate between them thus elongating their recorded durations. The solution was to use the standard deviation of the mean threshold which thresholds at a higher value making it easier to differentiate between the two bursts however this caused the bandwidths to shorten.

The determined burst properties showed that the majority of the frequency bandwidths are hovering around 50MHz with a mean value of 63MHz, the duration of the bursts detected was approximately 5-7s, the drift rate mean value was 9 MHz/s, and the burst velocity is approximately 0.3c. These values matched well with the commonly accepted values of wide bandwidths, 10 MHz/s and 0.3c respectively. However the duration was slightly high by a few seconds, more refining is required to the threshold and transform application to rectify this, but a possible cause is the overlapping continuums.

References

Lobzin, V. V., I. H. Cairns, P. A. Robinson, G. Steward, and G. Patterson (2009), Automatic recognition of type III solar radio bursts: Automated Radio Burst Identification System method and first observations, *Space Weather*, 7

Reid, H. A. S., & Ratcliffe, H. (2014), *Research in Astronomy and Astrophysics*, 14, 773

M. White, Stephen. (2007)] *Solar Radio Bursts and Space Weather*. *Asian J. Phys.* 16

P. Dbrowski, Bartosz & Krancowski, Andrzej & Baszkiewicz, Leszek & Rothkaehl, Hanna. (2016). Prospects for Solar and Space Weather Research with Polish Part of the LOFAR Telescope. *Acta Geophysica*. 64

Benz, A. O.: 2002, *Plasma Astrophysics. Kinetic Processes in Solar and Stellar Coronae*. Dordrecht: Kluwer Academic Publishers

Bonnin, X & Aboudarham, J & Fuller (2011). Automated detection and tracking of solar and heliospheric features in the frame of the European project HELIO

Solar Data Analysis Center, NASA Goddard Space Flight Center, <<http://www.solarmonitor.org>>