IUCAA NCRA

RADIO ASTRONOMY WINTER SCHOOL 2020

ANALYSIS ON CORRUPTION DUE TO RFI FROM GMRT DATA

NCRA • TIFR

Submitted by

Group 10

Aakash Narayan

Harsh Mishra

Vijayalakshmi V

Vysakh P A

Koidala Surya Prakash*

ABSTRACT

This experiment was conducted as a part of Radio Astronomy Winter School (RAWS) 2020 to learn about signals and the corruption due to RFI. When we observe the sky using a telescope, although we aim to receive the signals only from celestial sources, what we actually get is a mix of terrestrial and celestial signals. Man-made signals that get mixed with the celestial signals received by radio telescopes are termed as Radio Frequency Interference (RFI). For this experiment we used two ascii data files at a time each containing a list of 4194304(1024*4096) points. These were voltages recorded from the dipoles on the antennas C11 and C12 of the GMRT. Each value was recorded every 2.5e-9 s which is the Nyquist sampling for the signal recorded with a bandwidth of 200 MHz. The antennas were tracking the source 3C286 in the sky when the data were recorded.

RADIO FREQUENCY INTERFERENCE

IMPORTANCE OF UNDERSTANDING RFI

The radio signals arriving on Earth from astronomical objects are extremely weak, millions of times weaker than the signals used by communication systems. Because the cosmic radio sources are so weak, they are easily masked by man-made interference. Possibly even worse than complete masking, weaker interfering signals can contaminate the data collected by radio telescopes, potentially leading astronomers to erroneous interpretations.

Radio frequencies are divided up into blocks, or bands, designated for different types of uses. These international frequency designations are designed to prevent one type of station from interfering with stations of another type. A number of frequency bands are allocated to radio astronomy. However, transmitters using frequencies near those assigned to radio astronomy can cause interference to radio telescopes. This occurs when the transmitter's output is unduly "broad," spilling over into the radio astronomy frequencies, or when the transmitter emits frequencies outside its intended range. Other interference arises because radio transmitters often unintentionally emit signals at multiples of their intended frequency.

As use of radio for devices such as cellular telephones, wireless computer networks, garage door openers, and a whole host of other uses continues to increase, the threats to radio astronomy from inadequately engineered transmitters increases. A prime threat comes from transmitters in orbiting Earth satellites, since those transmitters are located overhead, precisely where radio astronomers must aim their telescopes to study the Universe.

AUTO-CORRELATION and CROSS-CORRELATION

Correlation is a mathematical tool used frequently in signal processing for analyzing functions or series of values, such as time domain signals (Wikipedia 2006). Correlation is the mutual relationship between two or more random variables (Ali). Autocorrelation is the correlation of a signal with itself (Parr 1999). This is unlike cross-correlation, which is the correlation of two different signals (Parr 1999).

Auto correlation function is a measure of similarity between a signal & its time delayed version. Consider a signals x(t). The auto correlation function of x(t) with its time delayed version is given by

$$R_{11}(\tau) = R(\tau) = \int_{-\infty}^{\infty} x(t)x^*(t-\tau)dt \quad [+ve\ shift]$$

$$= \int_{-\infty}^{\infty} x(t)x^*(t+\tau)dt \quad [-ve\ shift]$$

Where τ = searching or scanning or delay parameter.

Power density spectrum can be calculated by using the formula:

$$P = \sum_{n=-\infty}^{\infty} |C_n|^2$$

Cross correlation is the measure of similarity between two different signals. Consider two signals $x_1(t)$ and $x_2(t)$. The cross correlation of these two signals $R_{12}(\tau)$ and $R_{12}(\tau)$ is given by

$$R_{12}(\tau) = \int_{-\infty}^{\infty} x_2(t)x_1^*(t-\tau)dt \quad [+ve\ shift]$$

$$= \int_{-\infty}^{\infty} x_2(t)x_1^*(t+\tau)dt \quad [-ve\ shift]$$

Where τ = searching or scanning or delay parameter.

Parseval's Theorem

Parseval's theorem for energy signals states that the total energy in a signal can be obtained by the spectrum of the signal as

$$E = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int |X(\omega)|^2\, d\omega$$

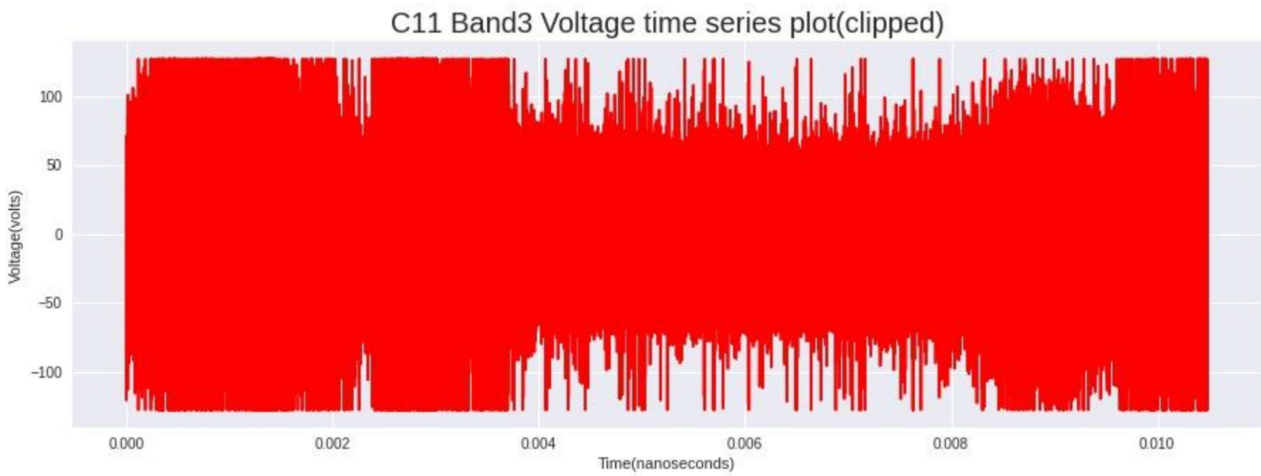
GMRT DATA ANALYSIS

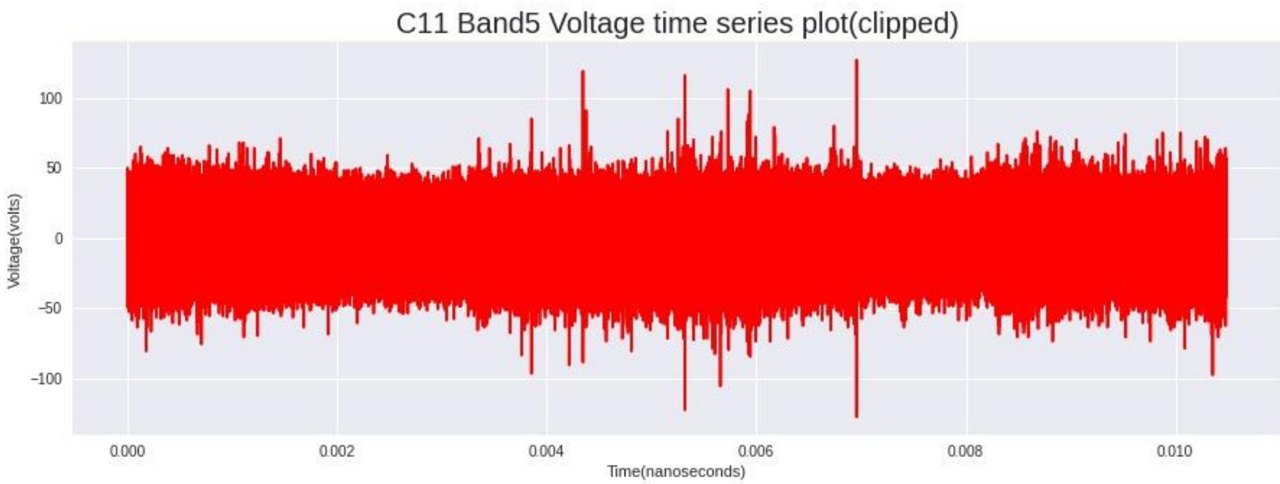
In this experiment we use the data recorded from the dipoles on the antennas C11 and C12 of the GMRT. The antennas were tracking the source 3C286 in the sky when the data were recorded. 3C 286, also known by its position as 1328+307 (B1950 coordinates) is a quasar at redshift 0.8493 with a radial velocity of 164,137 km/s. It is part of the Third Cambridge Catalogue of Radio Sources.

3C 286 is one of four primary calibrators used by the Very Large Array (along with 3C 48, 3C 138, and 3C 147). Visibilities of all other sources are calibrated using observed visibilities of one of these four calibrators.

The data was collected and stored in the form of ascii files containing a total of 4194304(1024*4096) points. Each correspond to voltage data measured across the dipoles of these antenna. Each value was recorded every 2.5e-9 s which is the Nyquist sampling for the signal recorded with a bandwidth of 200 MHz.

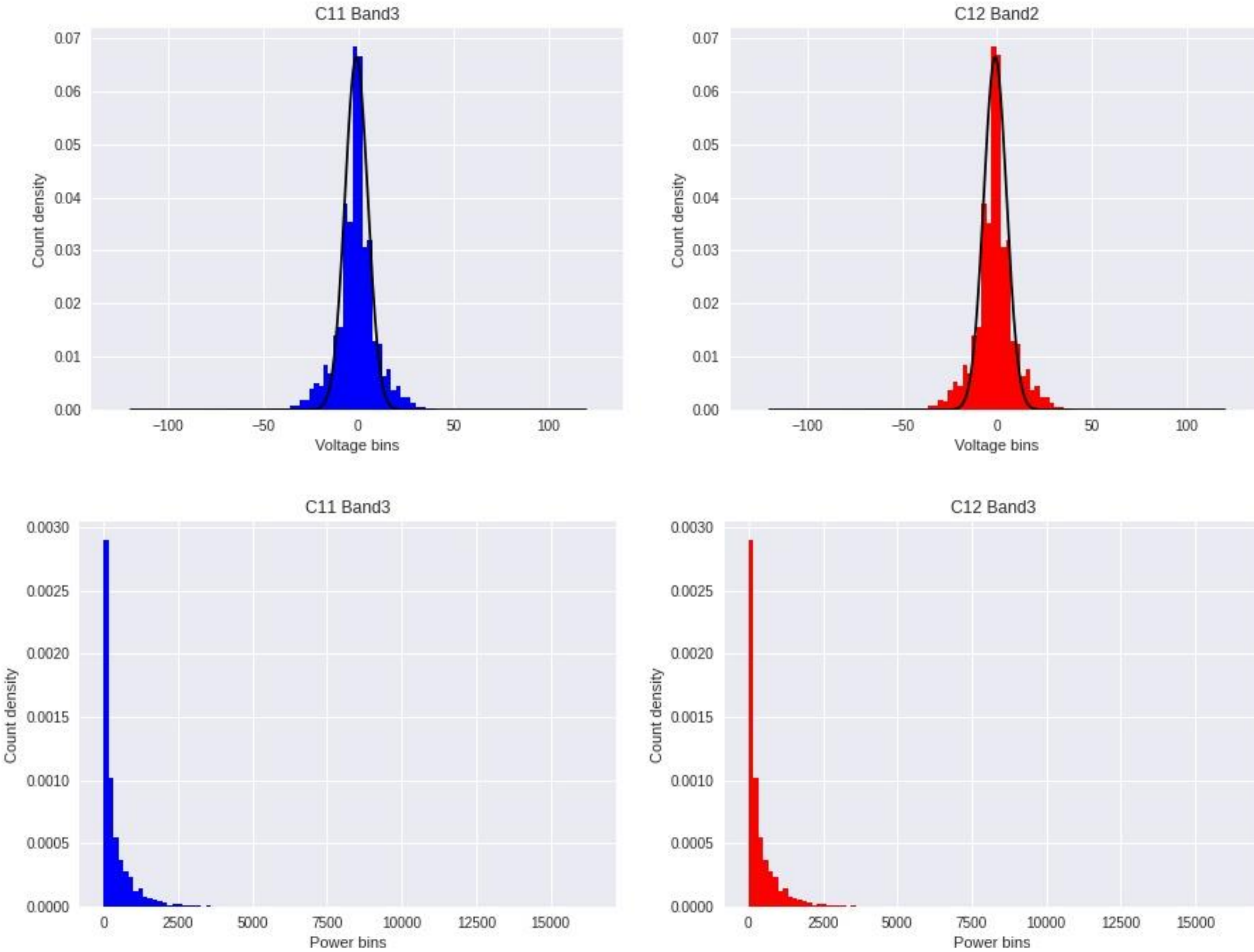
Two different data set is available, characterized by the band used. There are uGMRT Band 3(300-500MHz) data files and uGMRT Band 5(1050-1450MHz) data files.



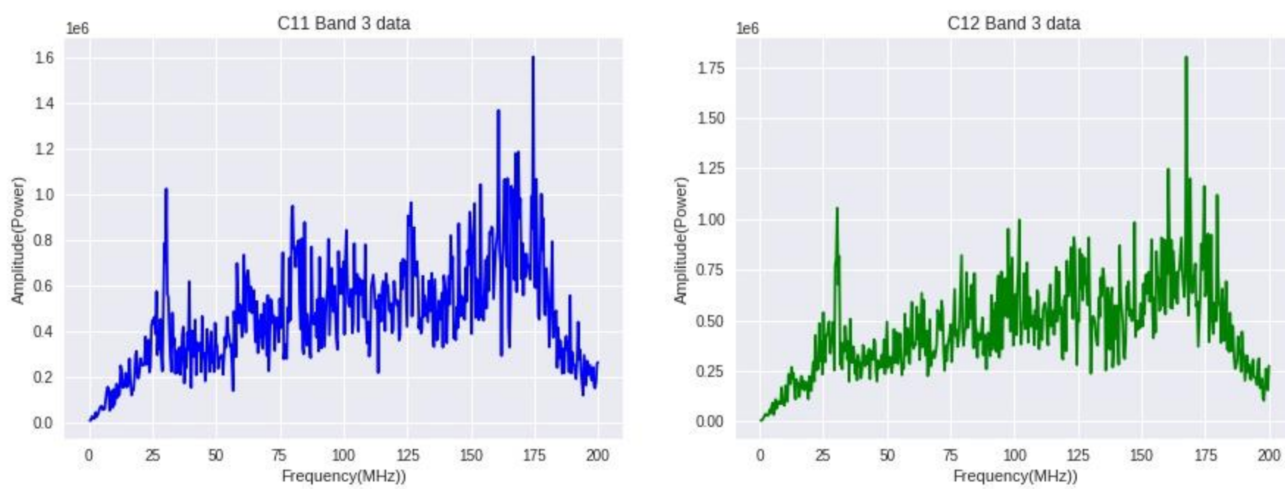


The First task was to plot the time series voltage data of the given signal. As you understand from the time series plot, the values of voltage varies from -120 to 120V for C11 band 3 data, after which the voltage is cutoff or saturation of the instrument is reached. In the case of Band5 data, the voltage tend to stay between - -100 to 100V itself.

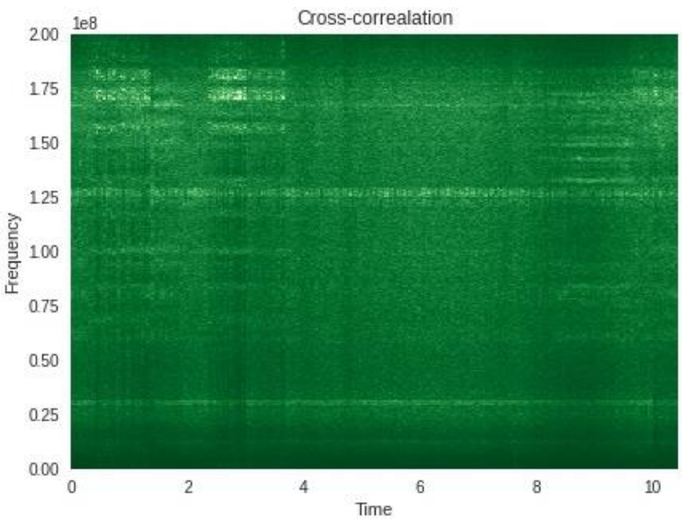
To further understand the data, we plotted the histograms for the voltage series. To remove the DC offset we subtracted the mean value from corresponding data to get the final data. We have also shared histogram of voltage square data.

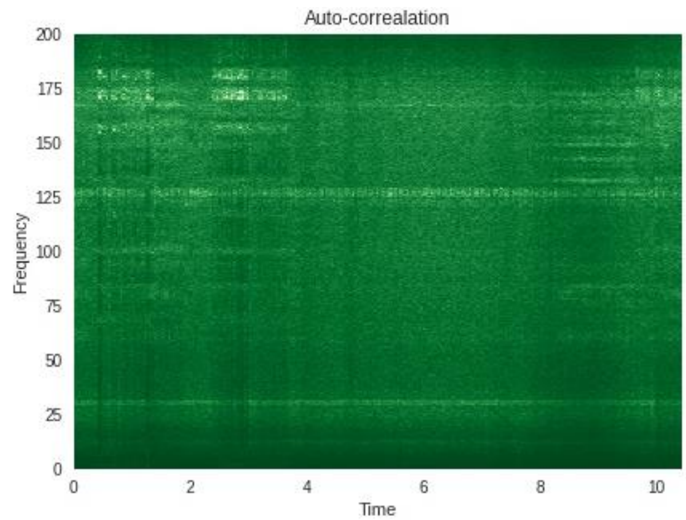


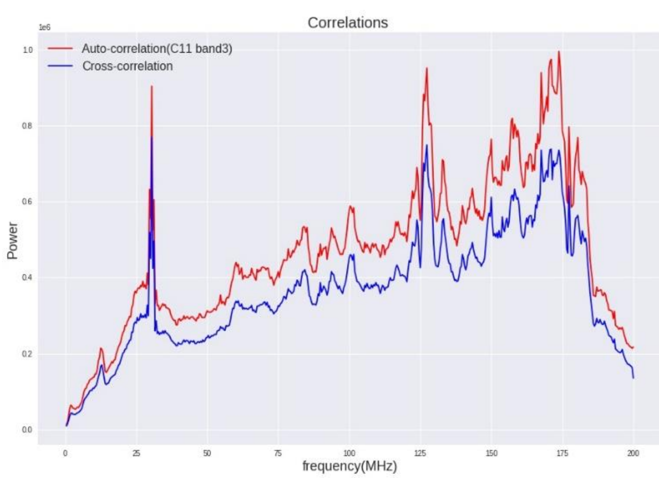
After performing the Fourier transform of these signals and plotting the amplitude vs Frequency plot, we are able to observe impulsive narrow peaks, which could imply the presence of background radio interference. Those plots are shown below.

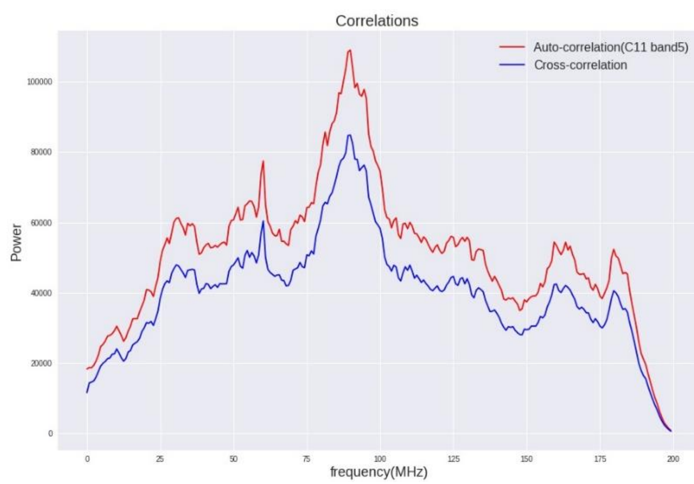


The next step was to preform auto-correlation and cross-correlation. We used both datasets from band 3 and band 5 to perform cross-correlation of the signals.





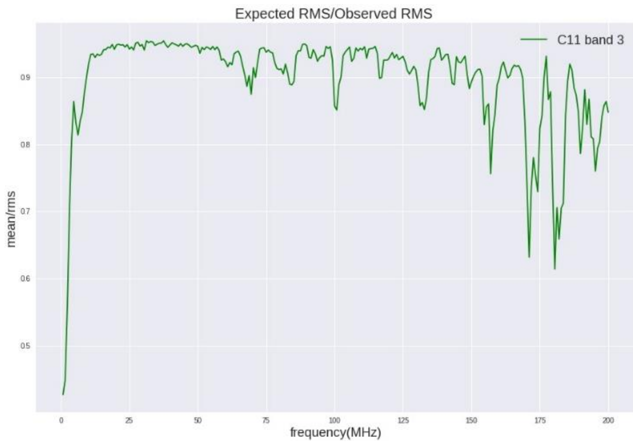


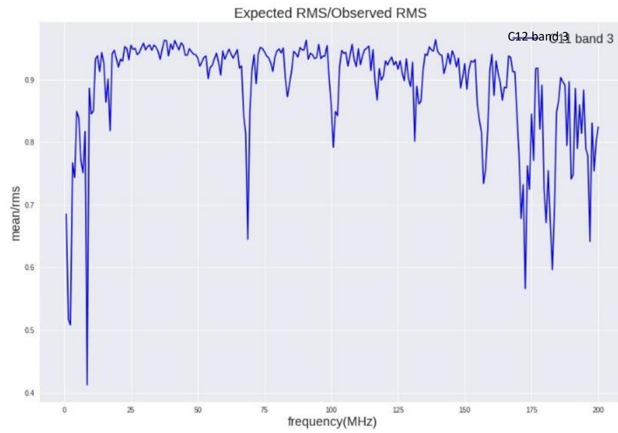


The plots shown below are the plot with frequency and time on X and Y axis and amplitude represented by the color scale.

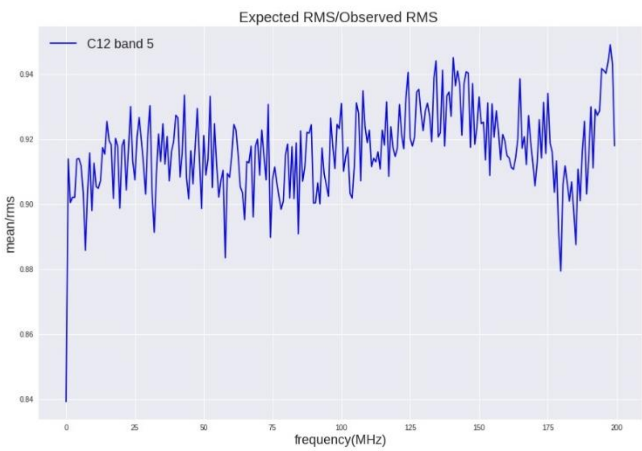
From the dynamic spectrum after auto correlation and cross correlation, we could see line which doesn't seem to change irrespective of the type of transformation. This could be concluded as due to radio frequency interference.

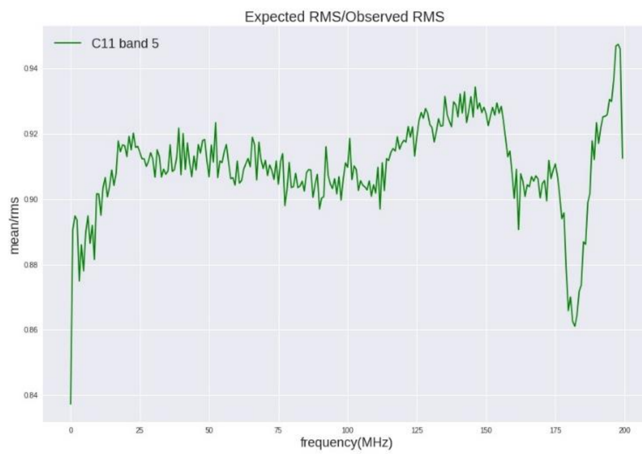
We could learn more about the RFI component of the spectra, by analyzing the statistical part of the data. We calculated the time-averaged mean and rms of the spectra. By taking the ratio we could learn more about the RFI presence in the signal. We could also use the ratio between expected and observed rms to find a more accurate measure to study about RFI.





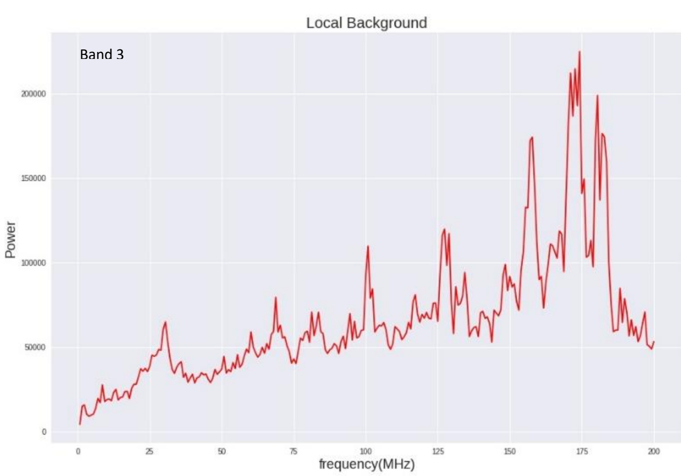
To compare the difference between different bands, we did the same with the band 5 data as shown below.

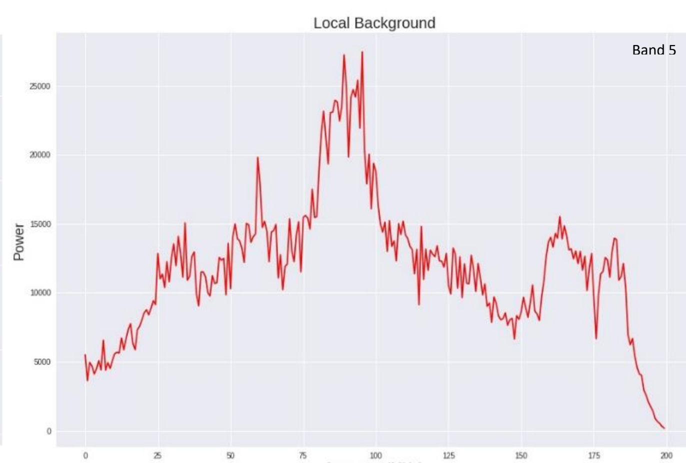




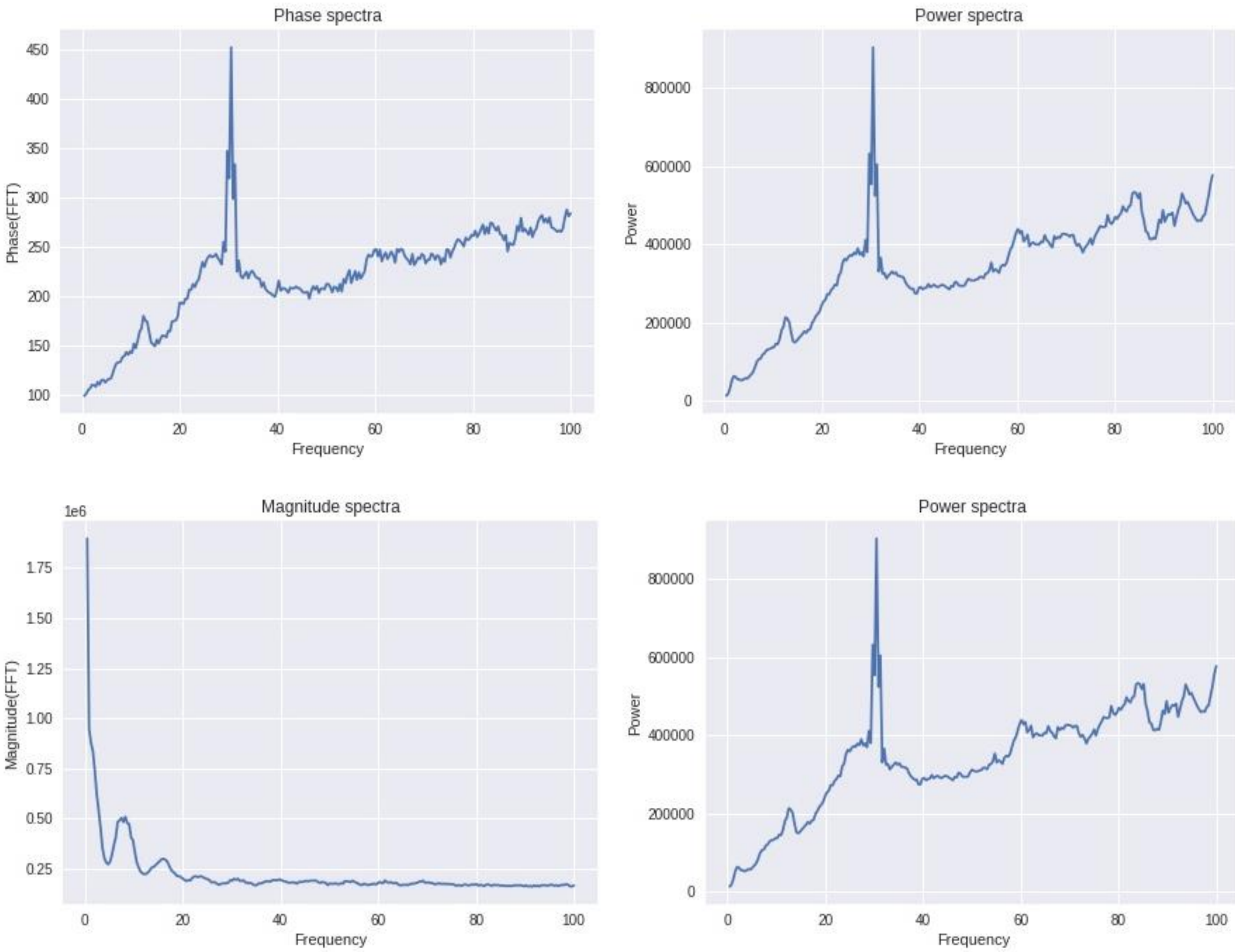
we can see here that larger band was more affected by RFI, this is expected as with a higher band, more frequencies are being accepted by the setup, thus allowing more unwanted noice in.

To get an idea about the local backgroud noise, we subtracted the geometric mean of the spectra from the cross spectra. we could the see the plot below.





One of the last task were to study the extend to which phase and magnitude of the voltage will affect the spectra. To study about phase, we removed the magniude part of the spectra, and plotted the dynamic spectra. we plotted the amplitude to frequency graphs, and compared it to the actual power spectra. the same procedure was follwed for the magnitude part to. Below are the results.



As you can see above in case of phase, it almost resembles the actual power spectra. So we can easily conclude that phase has high contribution towards the power spectra. On the contrary, magnitude shows the least similarity compared to the actual power spectra.

CONCLUSION

With all these methods we have understood the presence of the RFI in the signal as well as how much it will affect the measurements. We have also found that with a higher band, we receive more RFI compared to what we did with a lower band. We also found out that the phase part of the signal contribute much towards the final spectra compared to the magnitude part.

We could continue this further, by employing different RFI methods and studying which shows the best results.