RFI | Characterisation and Excision

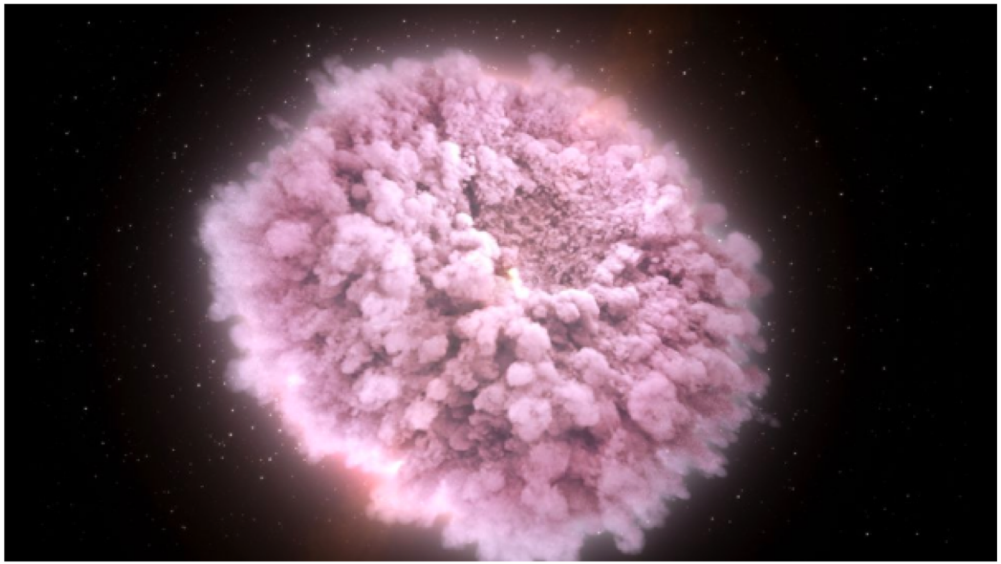
DARA 2021

By
Nadeem Oozeer
Data Science
RARG

Mysterious Gravitational Wave Sparks Days-Long Hunt — But It Was Just a Glitch

By Rafi Letzter - Staff Writer 3 days ago Space

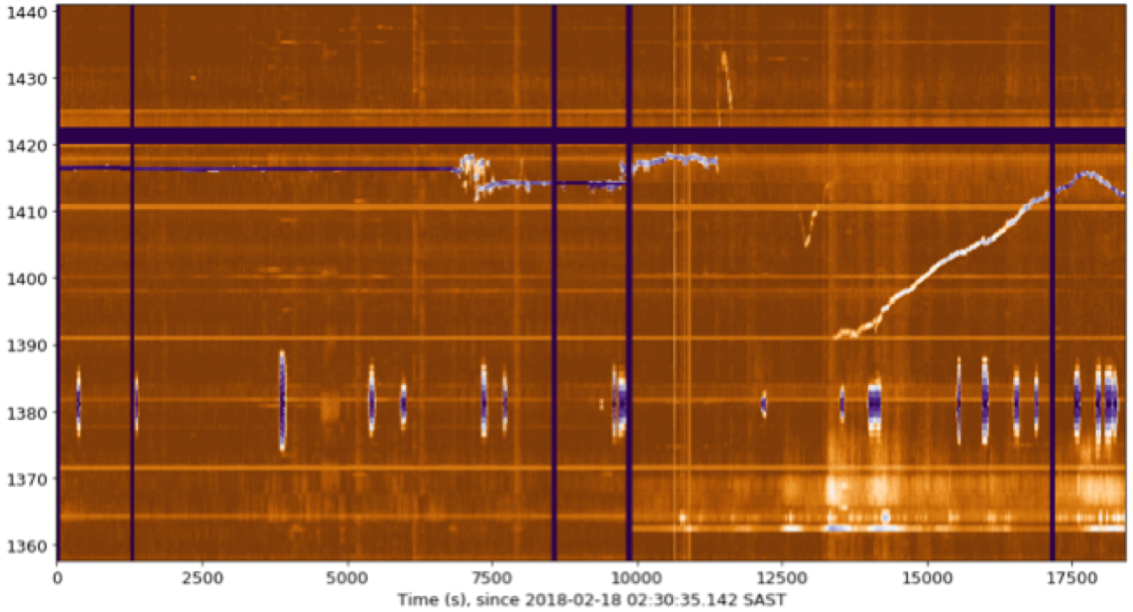
Astronomers and physicists all over the world were disappointed.



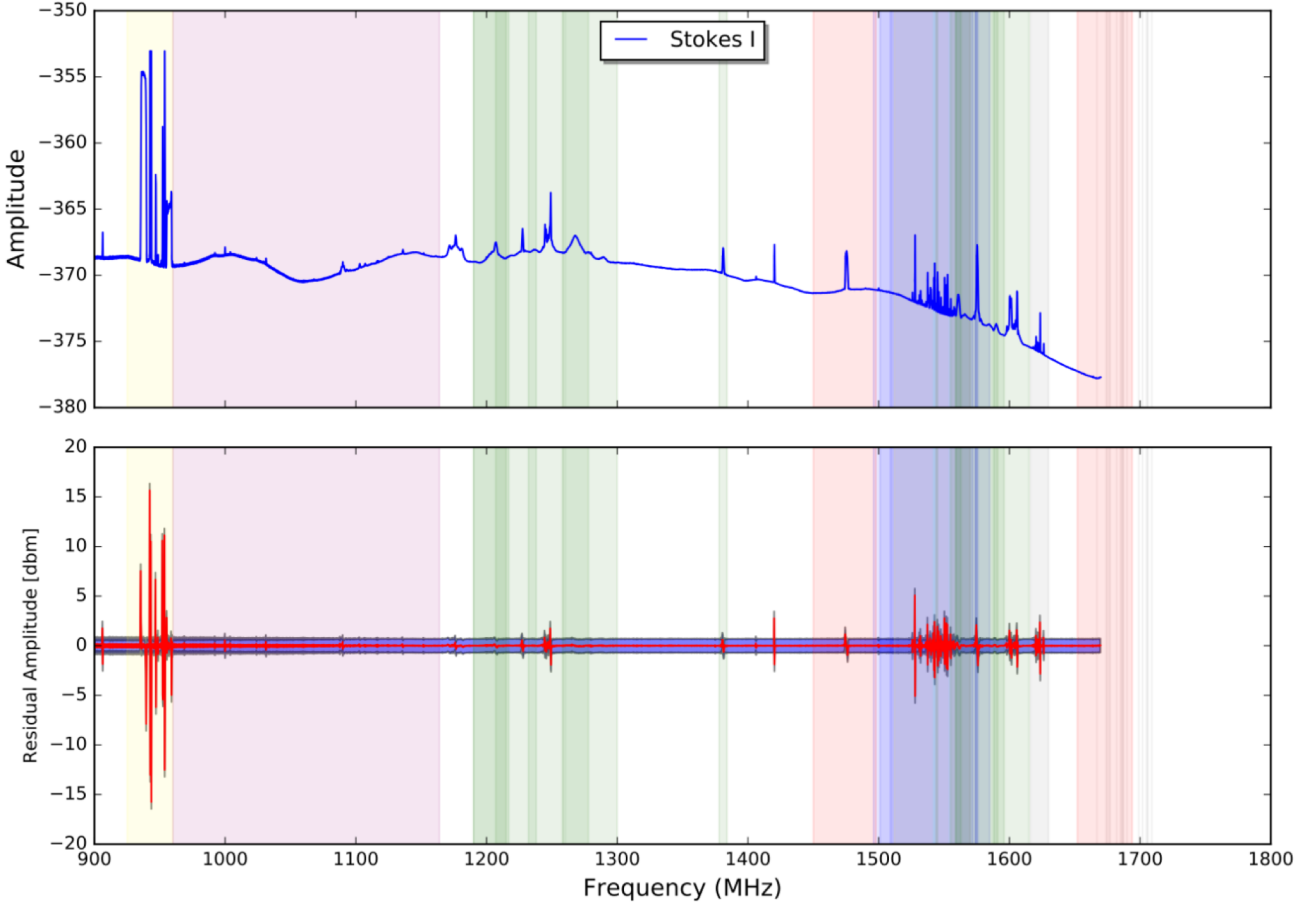
An illustration shows a supernova, one of the events raised as a possible source for the wave.
(Image: © NASA Goddard Space Flight Center)

Background

- RFI comes in various forms.
- RFI can for example be caused by television stations, aeroplanes and radar, while broadband RFI caused by phenomena such as lightning, high-voltage power cables and sparking electrical fences are often local in time.
- Strong RFI that is problematic is often either local in frequency or in time.
- A different class of RFI is caused by weakly transmitting but stationary -- and therefore systematic -- devices on site.



L-band RFI



Observer: RTS

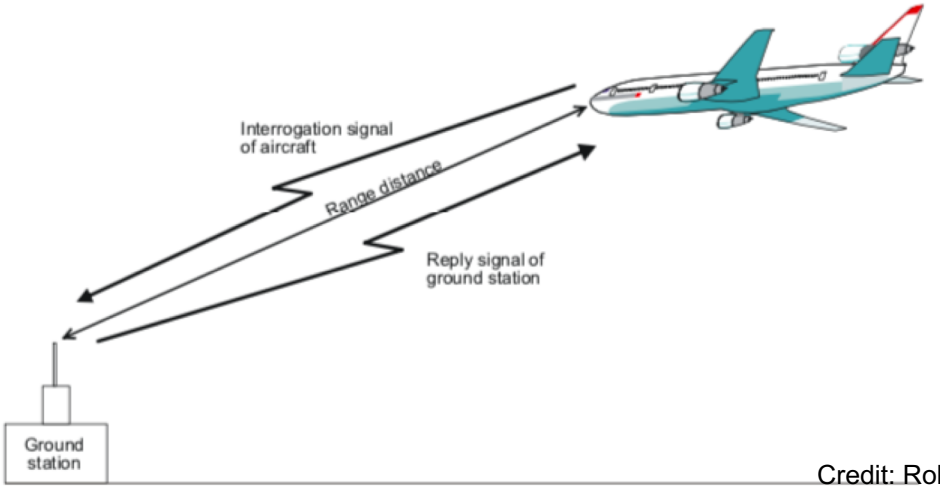
Product Type: 2.9 Self_Generated_RFI
Start time: 2015-08-04 16:25:14.874

Observation data:
/var/kat/archive/data/RTS/telescope_products/2015/08/04
/1438705513.h5

Antenna used: m063

Number of tracks:
111

Colour band used
Yellow - GSM
Purple - Astro Nav
Red - Afristar
Green - GPS sat



Credit: Rohde & Schwarz

DME RFI

Nadeem Oozeer, SARAO Data Science Team

Email : nadeem@ska.ac.za

B. Bassett, C. Finlay, I. Sihlangu

National Research Foundation

SARAO South African Radio Astronomy Observatory

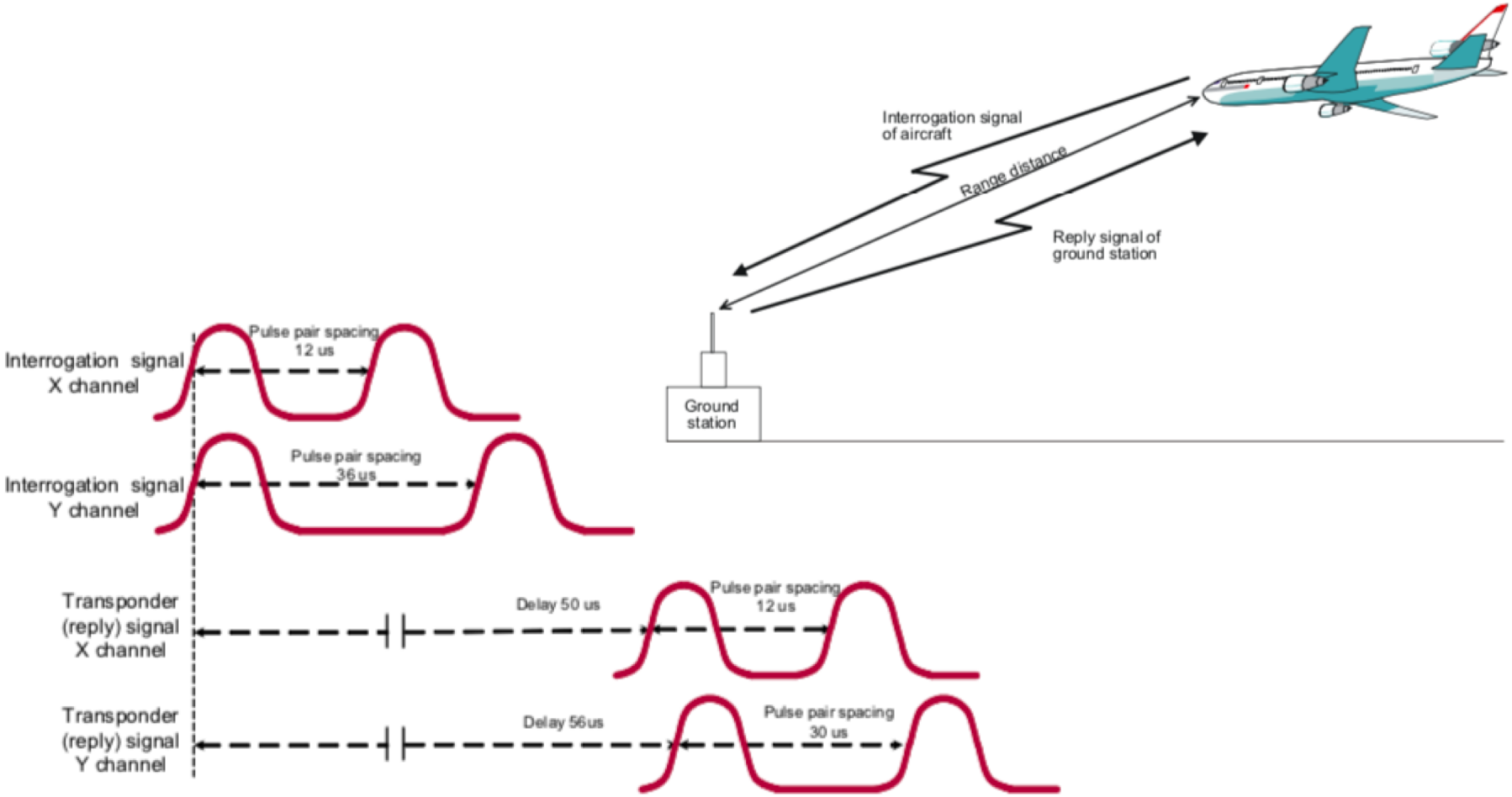
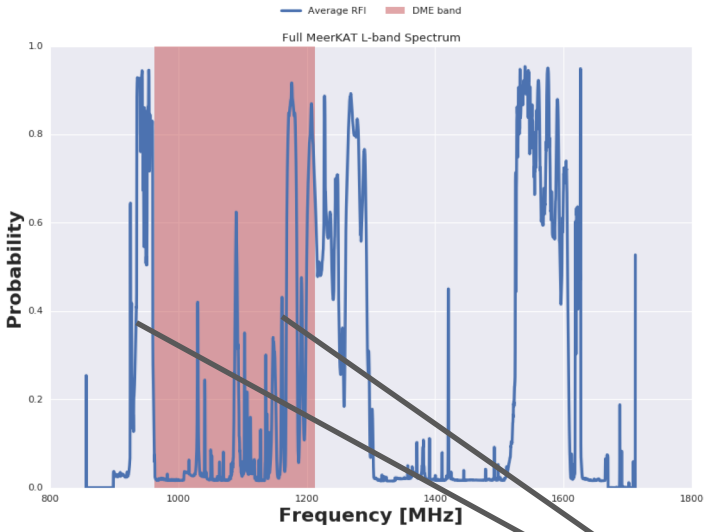
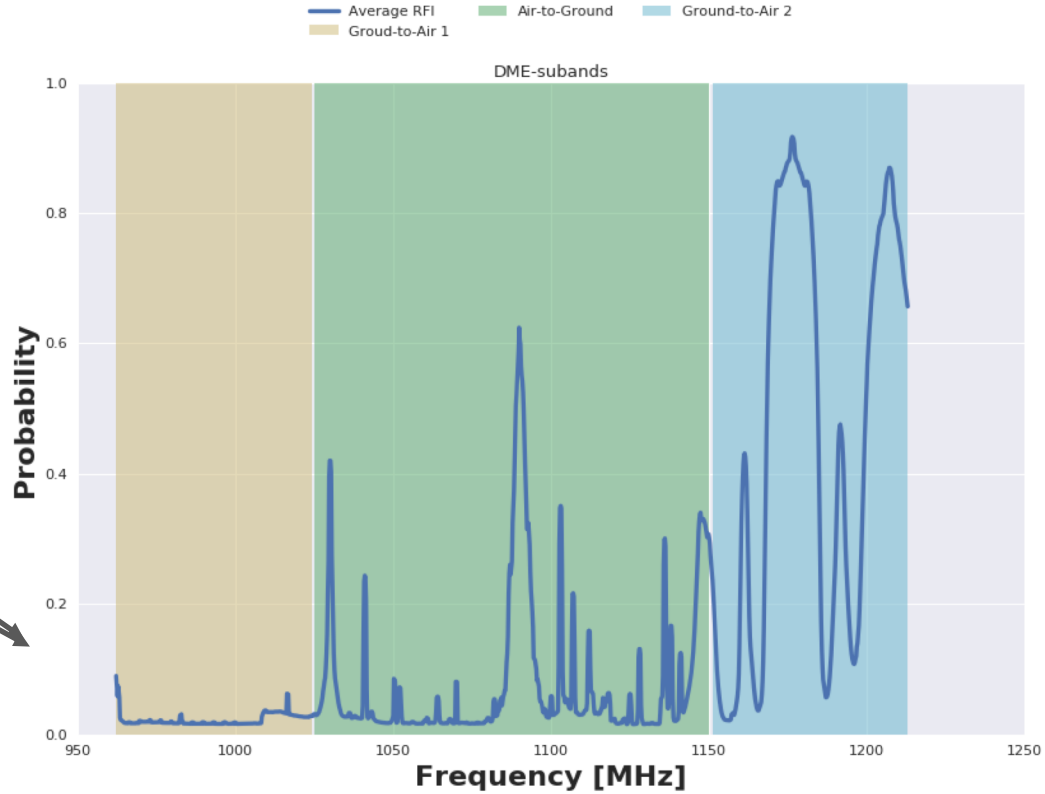
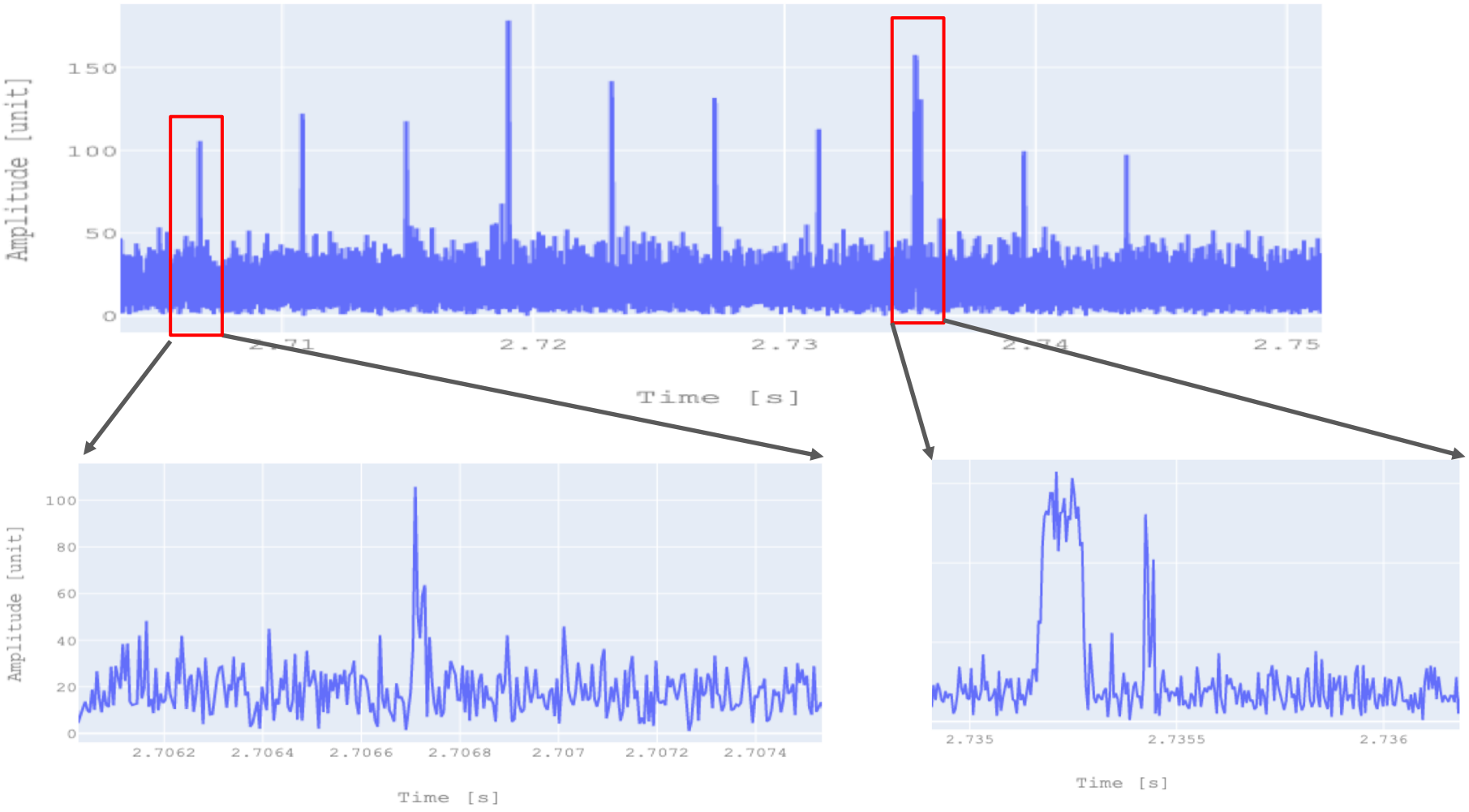


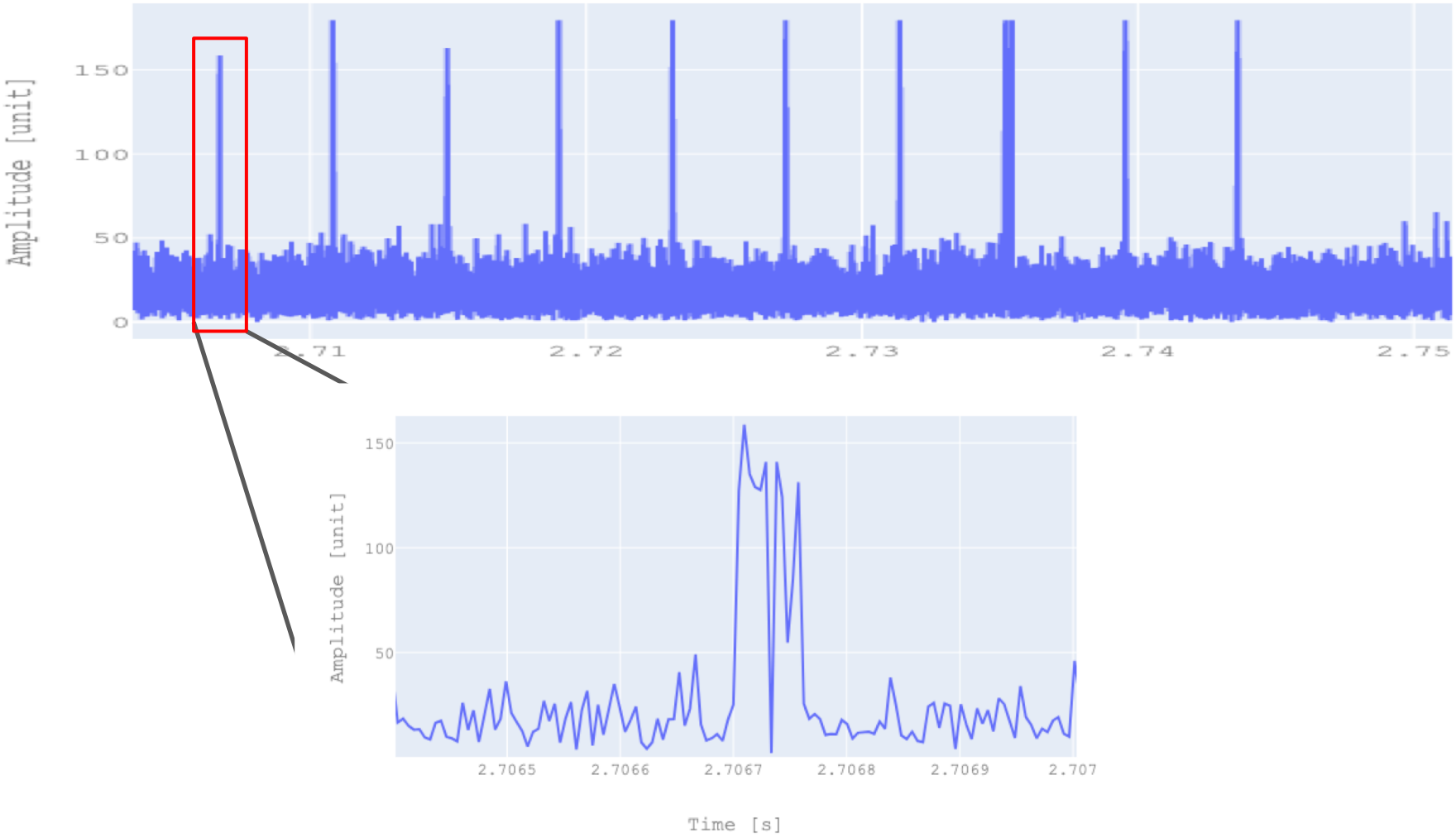
Figure 2: Time characteristic of DME signal envelope for X and Y channel.

Credit: Rohde & Schwarz


Average RFI
DME band
Full MeerKAT L-band Spectrum
Probability
Frequency [MHz]
0.0
0.2
0.4
0.6
0.8
1.0
800
1000
1200
1400
1600
1800


Average RFI
Air-to-Ground
Ground-to-Air 2
Groud-to-Air 1
DME-subands
Probability
Frequency [MHz]
0.0
0.2
0.4
0.6
0.8
1.0
950
1000
1050
1100
1150
1200
1250


Amplitude [unit]
150
100
50
0
2.71
2.72
2.73
2.74
2.75
Time [s]
Amplitude [unit]
100
80
60
40
20
0
2.7062
2.7064
2.7066
2.7068
2.707
2.7072
2.7074
Time [s]
2.735
2.7355
2.736
Time [s]


Amplitude [unit]
150
100
50
0
2.71
2.72
2.73
2.74
2.75
Amplitude [unit]
150
100
50
2.7065
2.7066
2.7067
2.7068
2.7069
2.707
Time [s]

MeerKAT RFI

Historical Probabilities (HP)

Isaac Sihlangu, SARAO Data Science Team

Email : isihlangu@ska.ac.za

R. Armstrong, B. Bassett, C. Finlay, N.Oozeer

Karoo

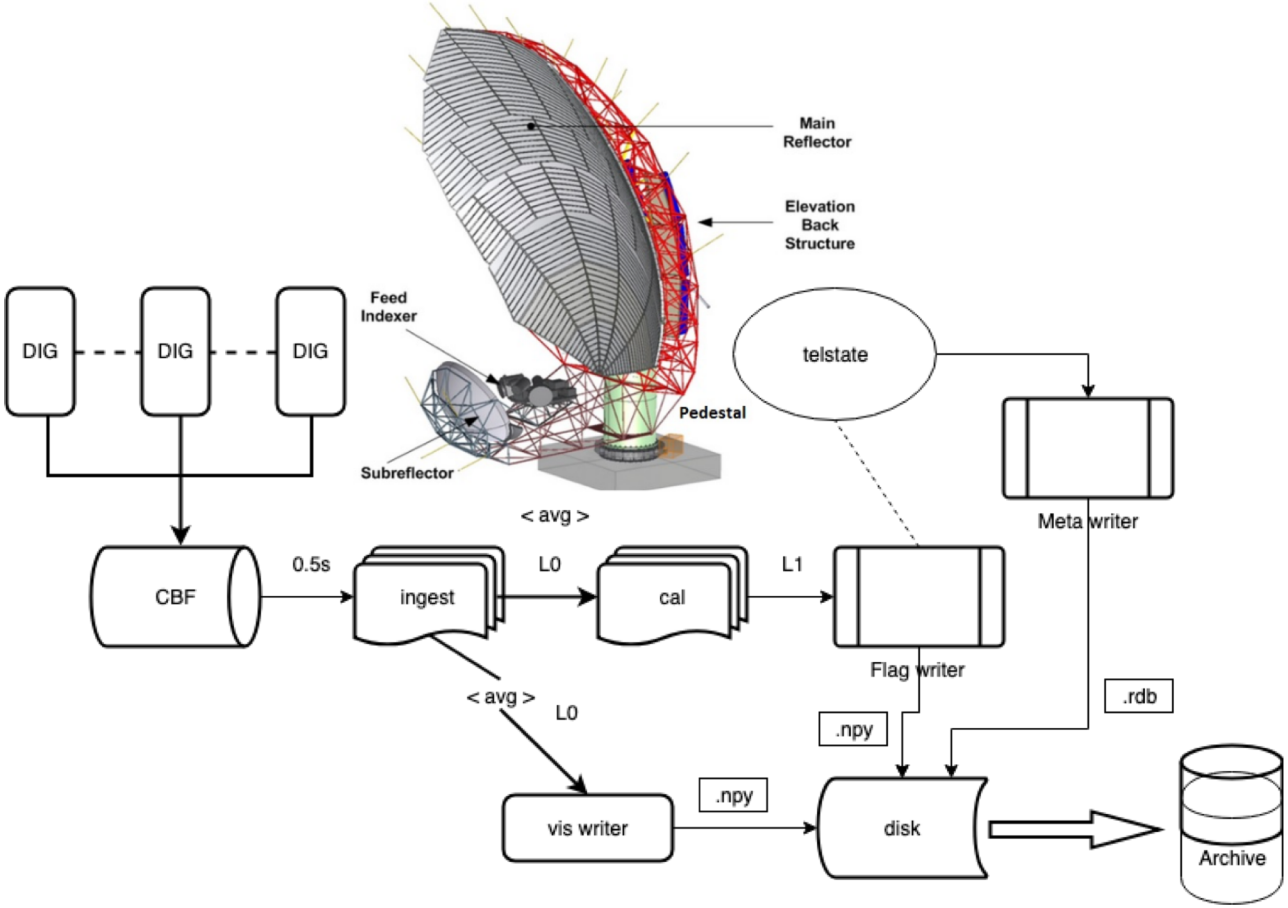
Array

Telescope

Historical

Probability

RFI



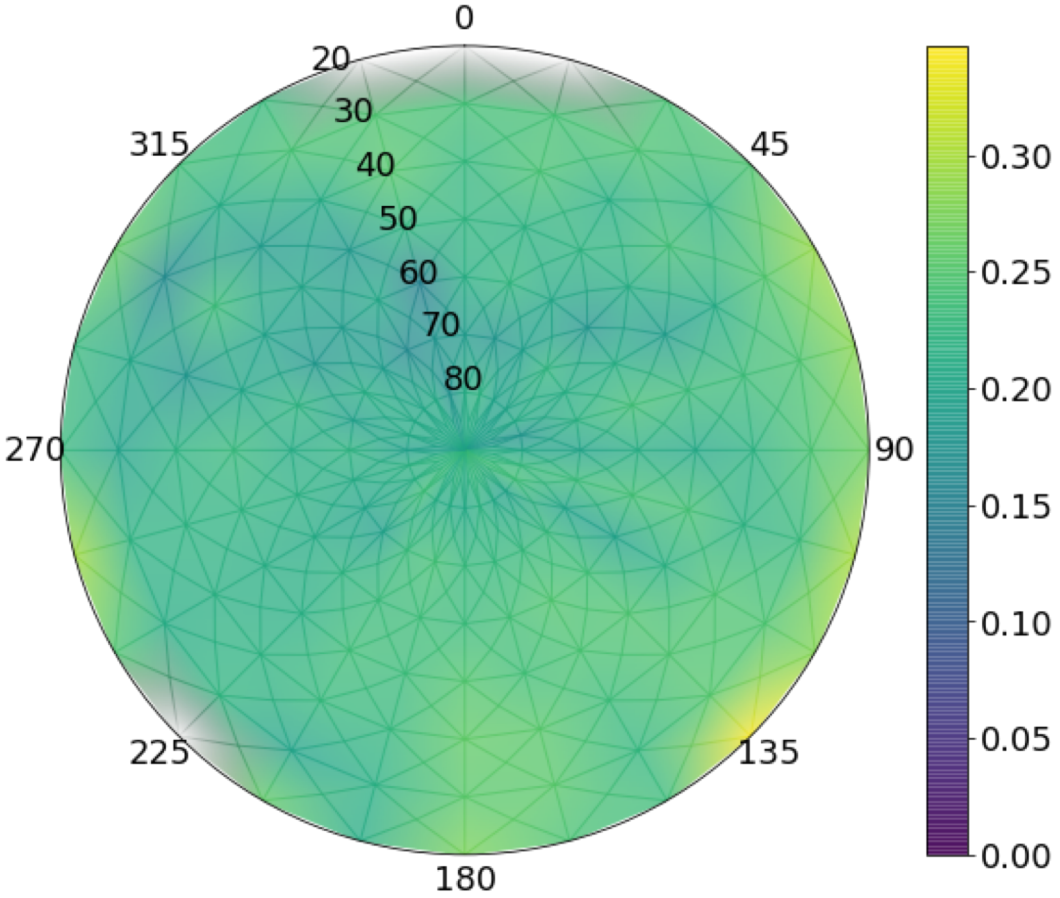
Isaac Sihlangu et al. https://arxiv.org/pdf/2008.08877

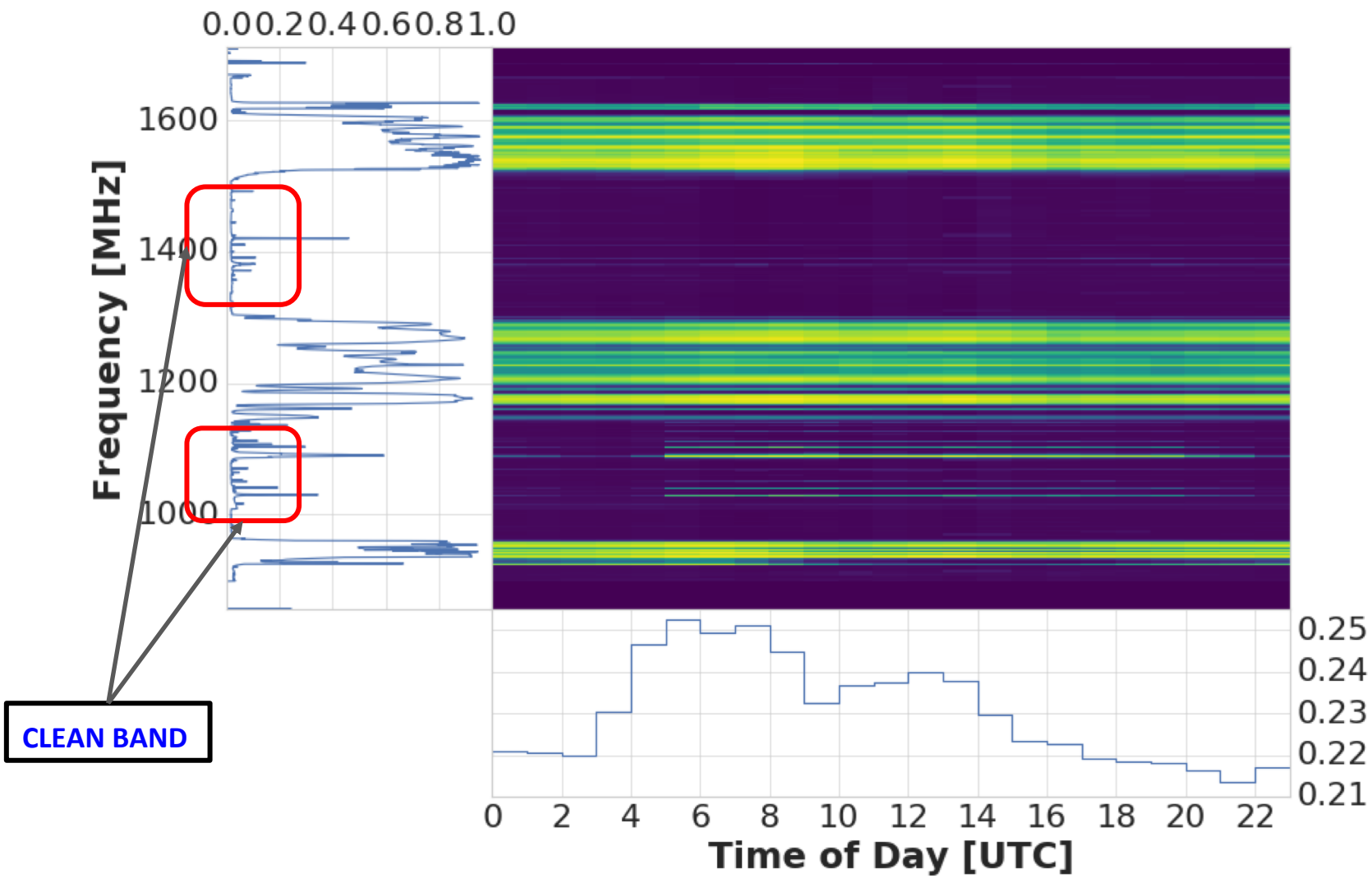
Use cases

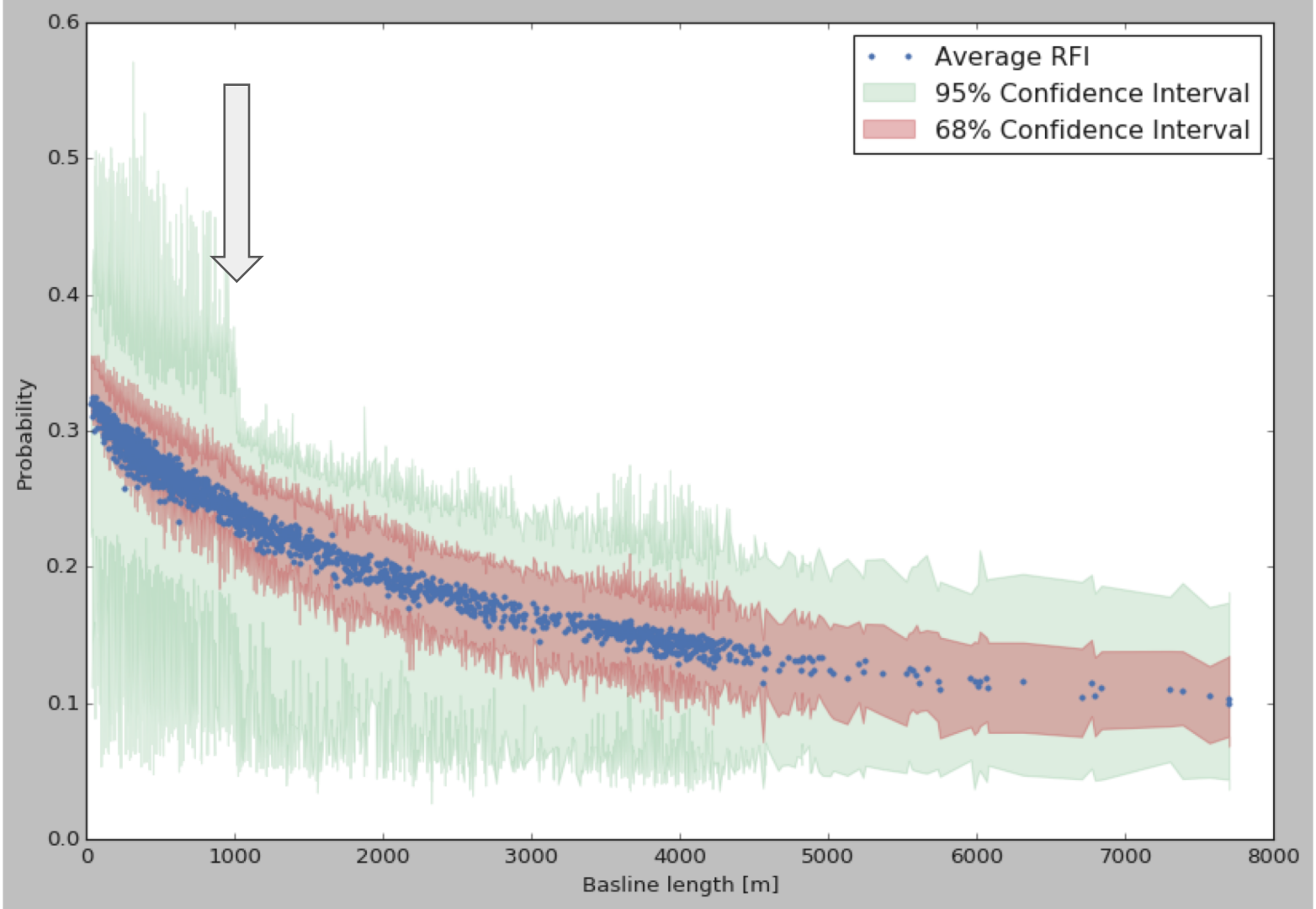
- **Operations:**
  - Schedule desired critical observations
- **RFI working group**
  - Provide an alert system if level of RFI at certain frequency goes beyond a critical threshold
- **Astronomers**
  - Optimise flagging
- **Policy makers**
  - Spectrum and telecoms team to monitor occupancy of dedicated frequency bands

Full sky - RFI probability

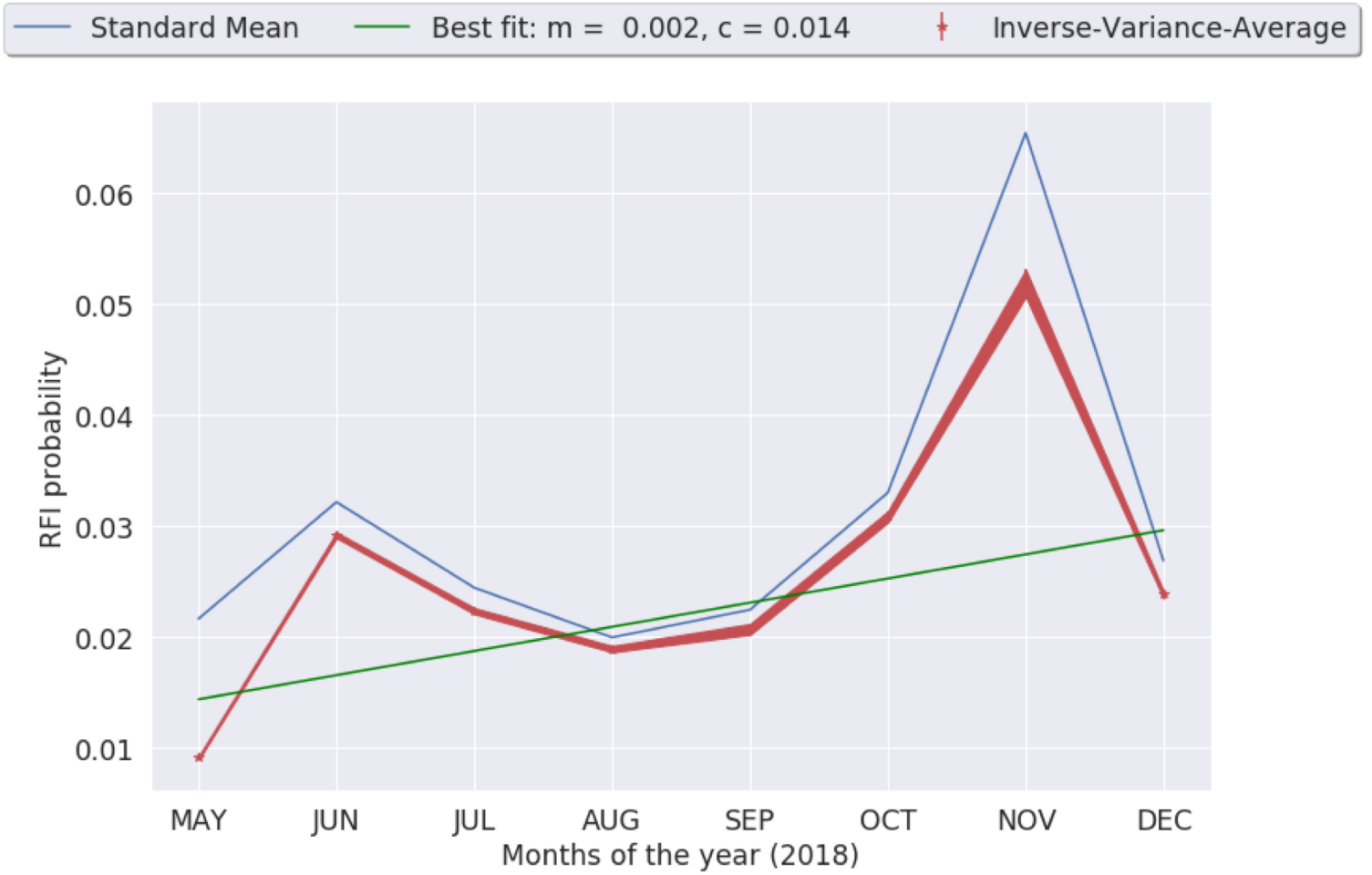
- May - Dec 2018
- 1500 hrs obs
- 400 TB imaging data
- T, F, B, Az, El

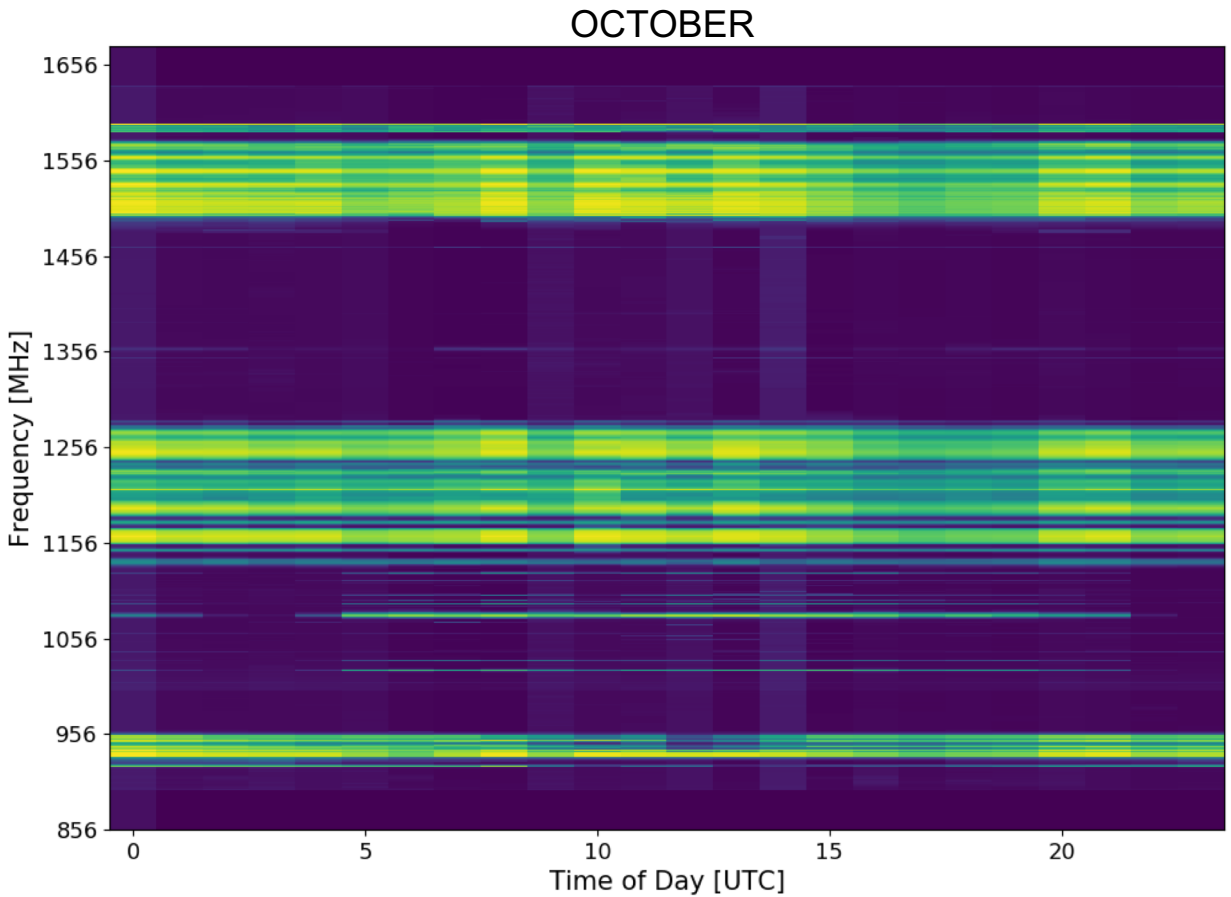


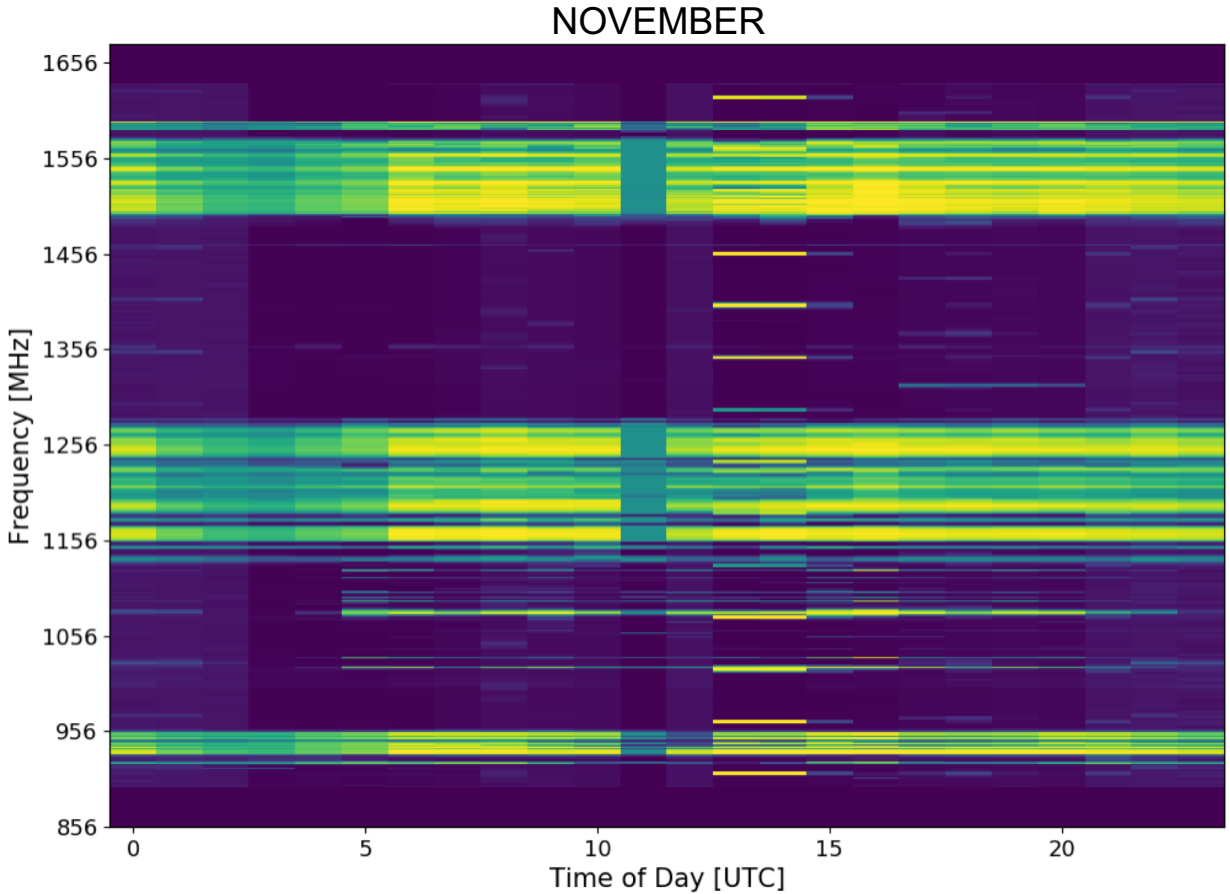

0.0 0.2 0.4 0.6 0.8 1.0
Frequency [MHz]
1600
1400
1200
1000
CLEAN BAND
0.25
0.24
0.23
0.22
0.21
0 2 4 6 8 10 12 14 16 18 20 22
Time of Day [UTC]

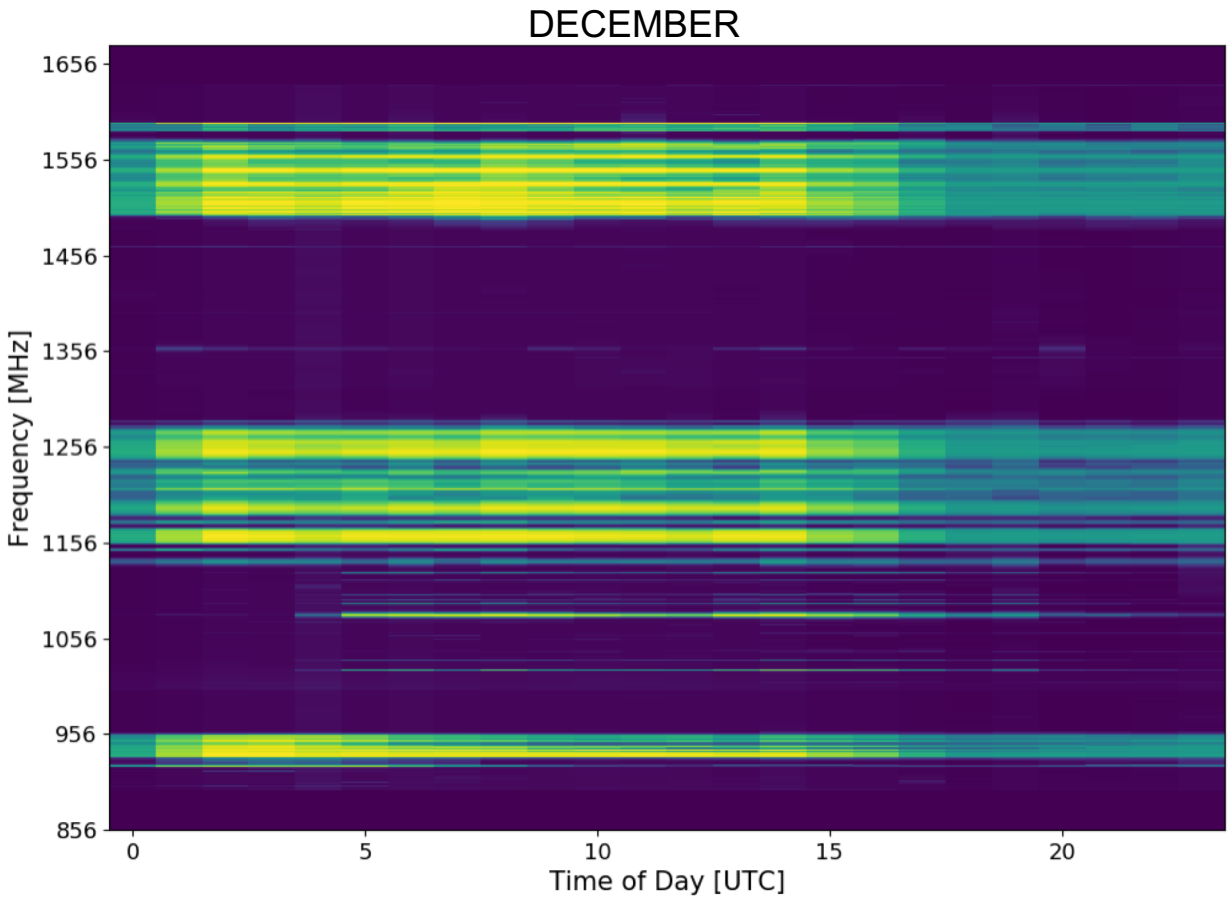

Average RFI
95% Confidence Interval
68% Confidence Interval
Probability
Basline length [m]
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
1000
2000
3000
4000
5000
6000
7000
8000

Time Evolution - RFI probability




OCTOBER
1656
1556
1456
1356
1256
1156
1056
956
856
Frequency [MHz]
0
5
10
15
20
Time of Day [UTC]


NOVEMBER
1656
1556
1456
1356
1256
1156
1056
956
856
Frequency [MHz]
0
5
10
15
20
Time of Day [UTC]


DECEMBER
1656
1556
1456
1356
1256
1156
1056
956
856
Frequency [MHz]
0
5
10
15
20
Time of Day [UTC]

Editing RFI

- Bad antennas -> remove them
- Correlator malfunctions -> remove time steps
- Initial pointing delay -> quack first timesteps
- Bandpass issue -> remove channels
- Low elevation -> elevation cut-off
- Interference -> remove antennas, timestep, frequencies or baselines..

Can **NOT** calibrate in the presence of issues in the data

Excising RFI

- Manual detection
  - Subjective
  - Enormous data volumes, computationally expensive…
- Automatic Thresholding
  - AoFlagger
  - RFIClean
  - Casa tools

RFI strategies

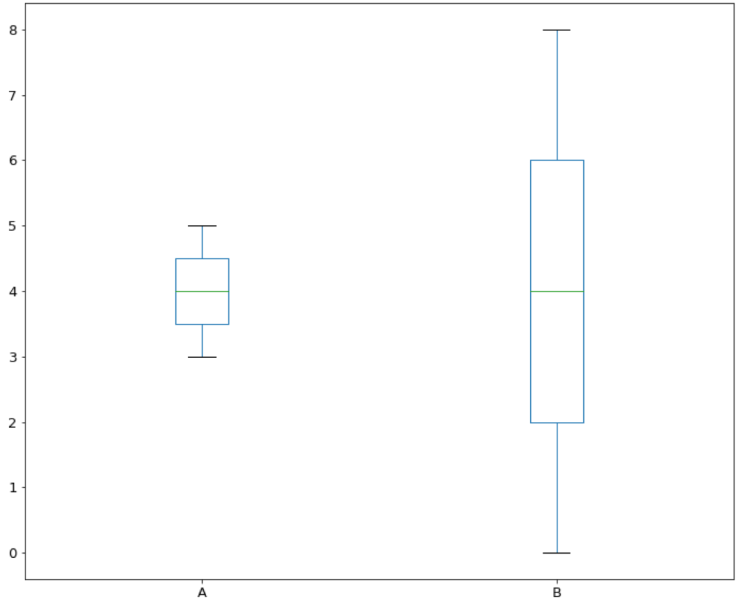
- Online pre-/post-correlation mitigation
  - Computationally expensive
  - Require high time resolution
- Offline
  - Post-optimized, data can be re-ordered and manipulated

Automated excision

- Detection
  - Find and throw away affected dat
  - Filtering or subtracting: estimate RFI contribution and restore affected data
- Detection (“flagging”)
  - Flag regions of bad data and do not use them in calibration and try correction afterwards
- Filtering
  - Hard, resulting data quality is not well understood, computationally expensive.

Mathematical tools

- Measures of spread
  - Mean, median, skewness and kurtosis
    - What should we use?
    - A = [3, 4, 5]
    - B = [0, 4, 8]
    - What are the mean?
    - What can be the issue with the Mean?



- Kurtosis

$$\beta_2 = \frac{\mu_4}{(\sigma^2)^2}$$

β_2 : Kurtosis.

μ_4 : Fourth moment of the probability distribution function.

σ^2 : Variance of the probability distribution function.

- Median Absolute deviation (MAD)

$$M = Median(X)$$

$$MAD = Median|X - M|$$

$$\sigma_r = 1.4826 * MAD$$

- Filtering:
  - MAD filter
  - sgolay...

MeerKAT

RFIsim

RFI Removal

Christopher Finlay, SARAO Data Science Team

Email :cfinlay@ska.ac.za

R. Armstrong, B. Bassett, C. Finlay, N.Oozeer

RFIsim

What do we want?

- A simulator that produces realistic visibilities corrupted by external RFI sources.
- To know the exact contribution of RFI in each visibility voxel.

Why do we want it?

- To compare different RFI detection and excision algorithms. Astronomers cannot provide what we need.

Where are we now?

A MeerKAT simulator that produces realistic RFI-corrupted visibilities.	
A dataset that is being used to train a neural network for RFI detection.	

BONUS

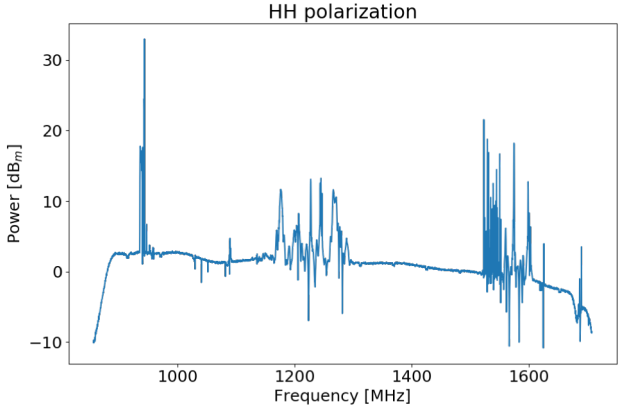
A way to predict satellite positions which has been used to optimize a MeerKAT early science observation proposal.

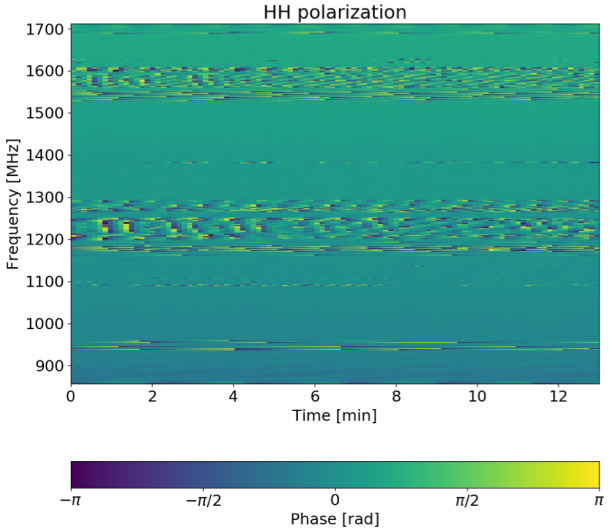
What effects are included?

Real satellite paths	✓
Ground-based RFI	✓
Realistic RFI characteristics (amplitude and frequency)	✓
Accurate astronomical sky model	✓
MeerKAT (u,v) coordinates	
Realistic Bandpass	✓
Realistic primary beam	✓
Time variation (gains, RFI amplitude)	✓

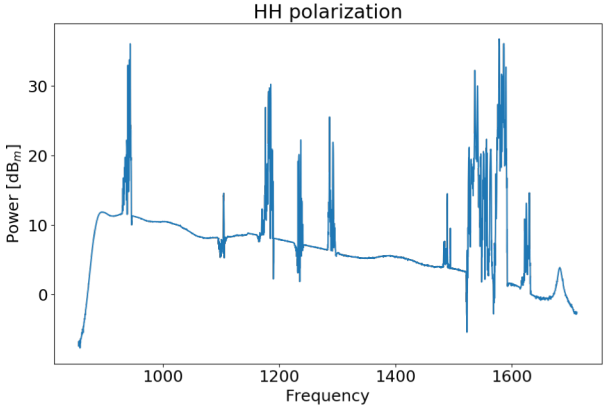
Results

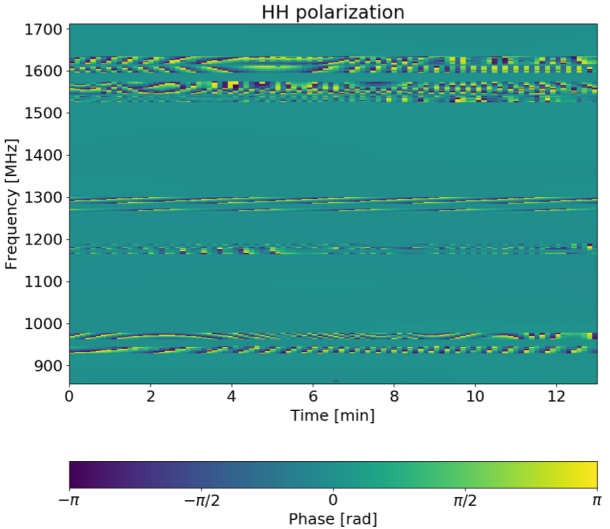
Real





Simulated





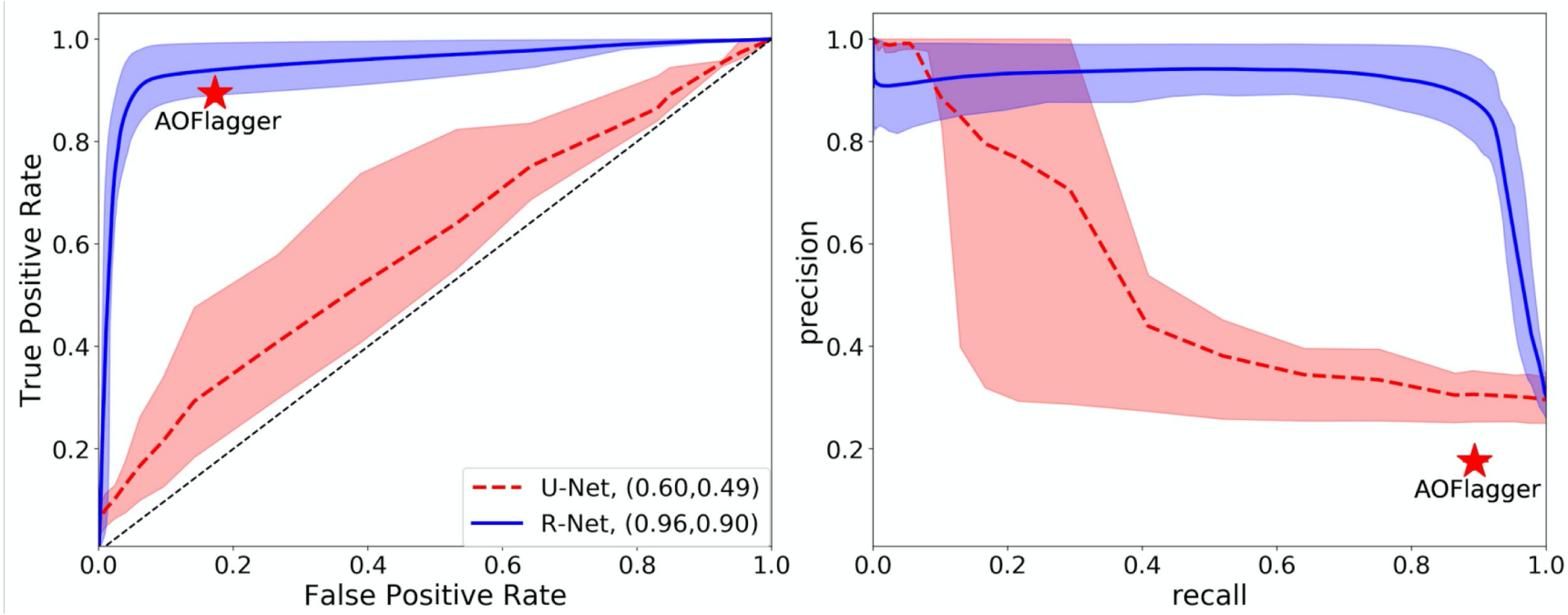
MeerKAT

RFI detection using **Deep Learning**

Alireza Vafaei Sadr, SARAO Data Science Team

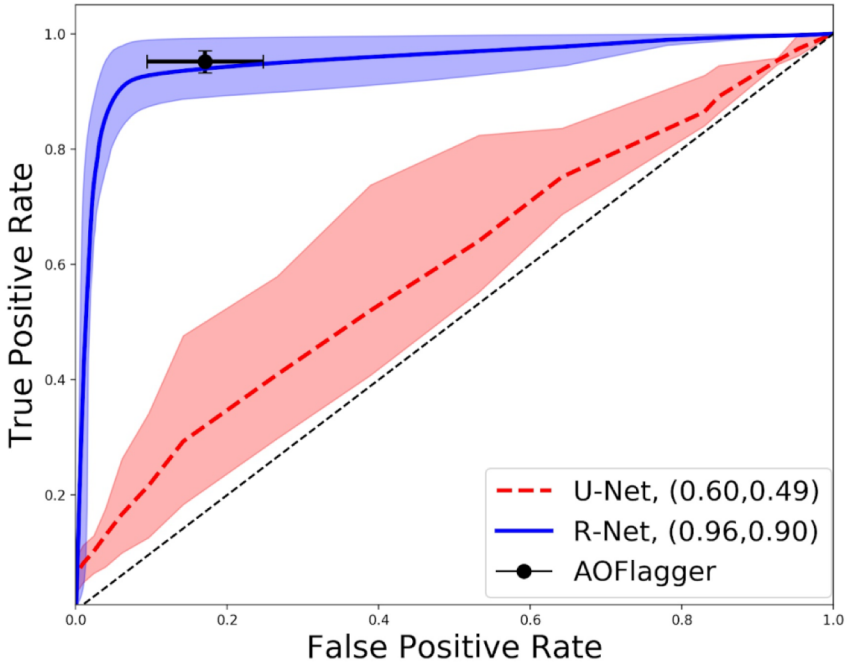
B. Bassett, (Y. Fantaye; C. Finlay), N.Oozeer, (R. Ramphul)

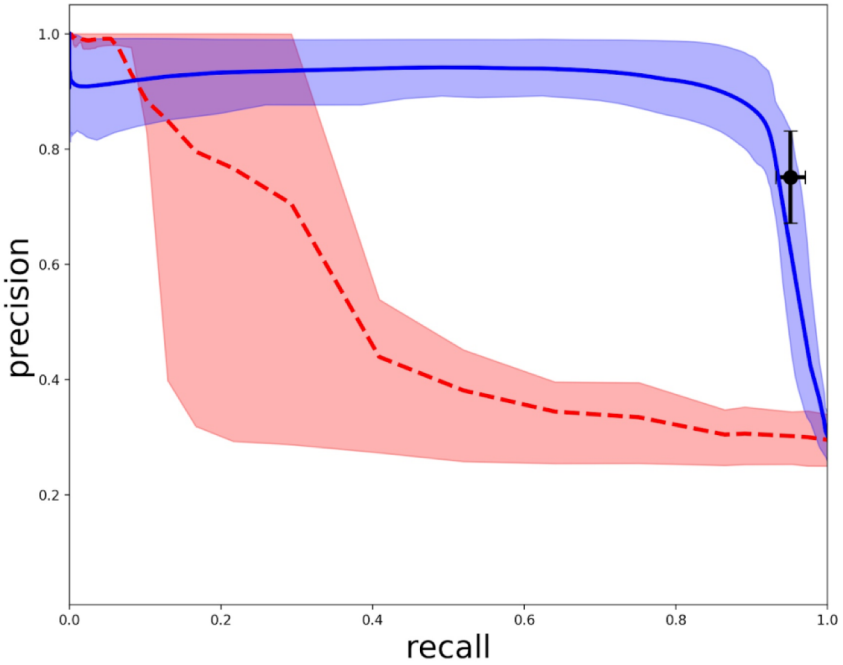
Comparison with Default AOflagger



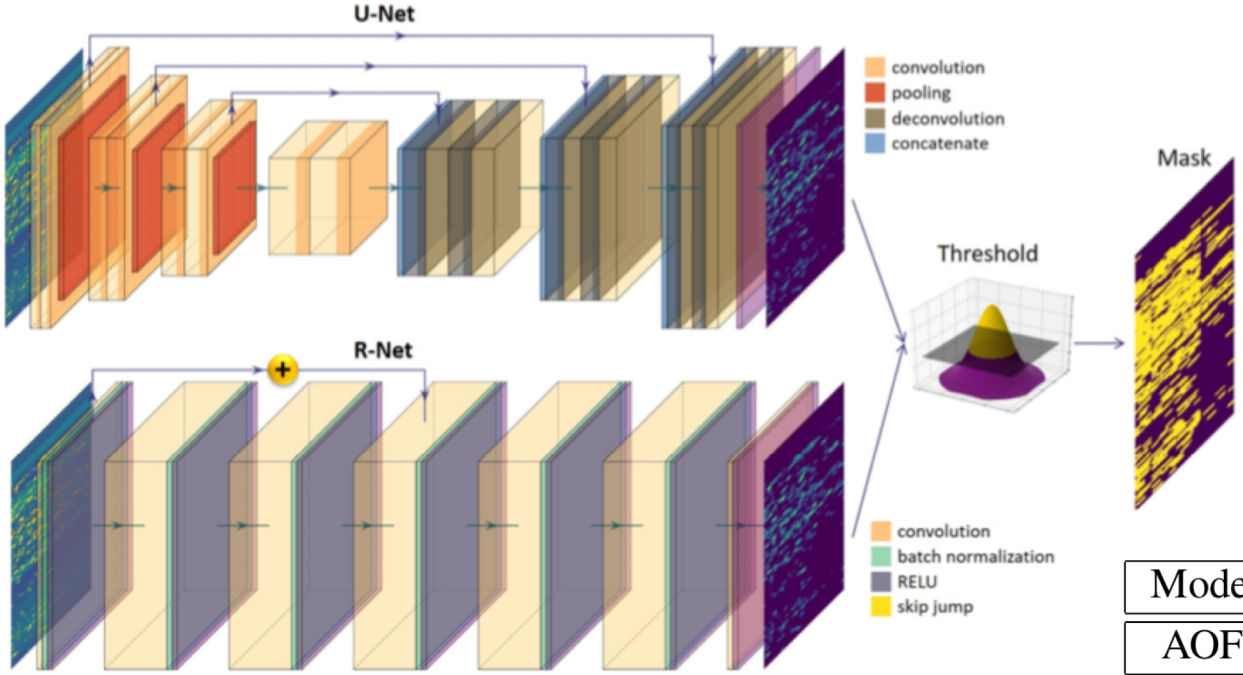
Deep learning is significantly better than AOFlagger unless AOFlagger is extensively tuned (which no one does). Our algorithm tunes automatically, and is flexible (not just a **single** False Positive Rate).

Comparison with Optimised AOflagger





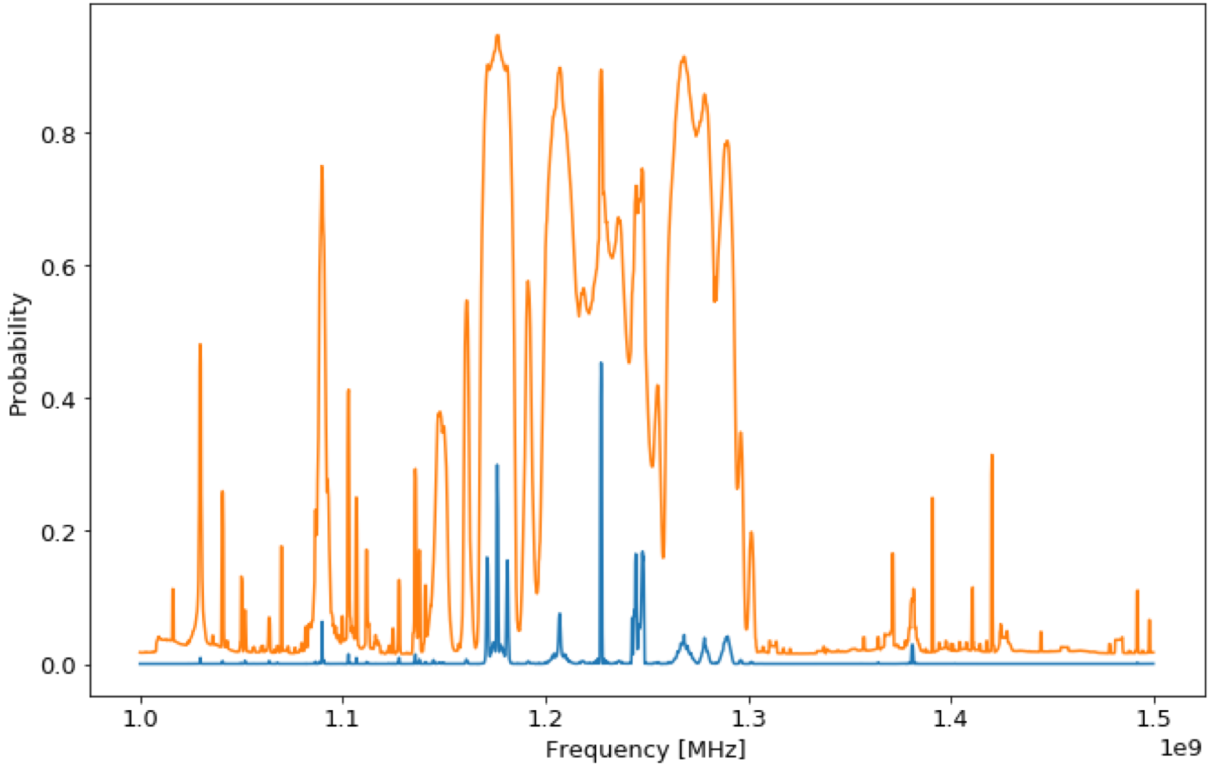
Network from KAT-7



Model	variables	exe. time(s)
AOFlagger	-	31
R-Net-5	11473	18
R-Net-6	15181	24
Unet-3-16	116770	49
Unet-3-32	465986	53

Frequency Band (MHz)	Shared Service	Scientific Observations
Frequencies below 100MHz		
25.550 - 25.670		RF Bursts (Sun, Jupiter), Interstellar Medium, Steep spectrum sources
37.5 - 38.25	Mobile Communication.	RF Bursts (Sun, Jupiter), Interstellar Medium, Steep spectrum sources
73.0 - 74.6		RF Bursts (Sun, Jupiter), Interstellar Medium, Steep spectrum sources
Frequencies below 1000MHz		
150.05 - 153.0	Mobile Communication except Aeronautical.	RF Bursts (Sun), Interstellar Medium, Steep spectrum sources, Pulsars
322.0 - 328.6	Mobile Communication.	Deutrium line (327.384 MHz),RF Bursts (Sun), Interstellar Medium, Steep spectrum sources, Pulsars
406.1 - 410.0	Mobile Communication except Aeronautical.	RF Bursts (Sun), Interstellar Medium, Steep spectrum sources, Pulsars
608.0 - 614.0	UHF TV Broadcasting.	RF Bursts (Sun), Interstellar Medium, Steep spectrum sources, Pulsars
Frequencies below 5000MHz		
1400.0 - 1427.0	Space Research and Earth Oration-Satellite. (passive)	Galactic and local extragalatic H1 (Hydrogen-1) (1420.406 MHz), Galactic Continuum
1610.6 - 1613.8	Mobile Satellite and Aeronautical Radio Navigation.	OH(Hydroxyl)
1660.5 - 1670.0	Mobile Satellite, Mobile Communication and Space Research.	OH(Hydroxyl)
1710.0 - 1930.0	Mobile Communication.	OH(Hydroxyl)
2655.0 - 2700.0	Mobile Communication except Aeronautical, Mobile Satellite, Space Research and Earth Exploration Satellite.	
4800.0 - 5000.0	Mobile Communication and Space Research.	Formaldehyde, Continuum, VLBI

Credit:N C Thompson, Thesis, 2014, Stellenbosch



Outlier detection from Ingest (blue) and flagger outlier(orange) for a RFI free region

- We noticed from the ingest detection that there is around a 17.5% chance of picking an outlier between frequency 1240 MHz - 1250 MHz.
- This corresponds to redshift z ~ 0.127 - 0.135.
- Any emissions from astronomical sources in this redshift range will inevitably be removed by the ingest and there is no way to undo the excision.

Conclusion

- Sources of RFI
- How to edit and excise RFI
- Why we need to excise RFI
- What are some of the tools to use
- New algorithm to detect RFI
- The impact of RFI on astronomy