Detecting Radio Frequency Interference (RFI)

in MeerKLASS Data

Tamera Kassie

Affiliations : SARAO, UWC (MSc)

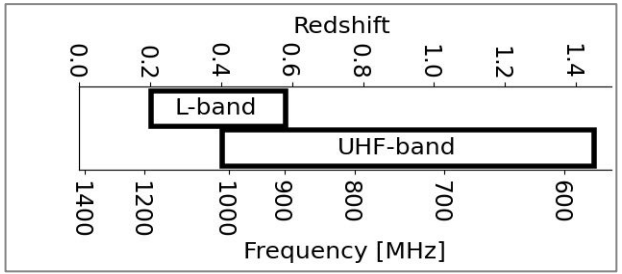
Supervisors: Prof. Mario Santos & Dr. Piyanat (Boom) Kittiwisit

AFAS Conference

24 March 2026

MeerKLASS : The MeerKAT Large Area Synoptic Sky survey

- MeerKLASS is an eXtra Large Survey Project (XLSP) of MeerKAT in UHF Band (500-1016 Mhz).
- MeerKLASS HI IM uses the MeerKAT radio telescope in single dish observing mode to create intensity maps (IM) of the southern sky.
- IM surveys, map the sky brightness over an integrated region. Challenged by foregrounds (synchrotron emission, RFI)
- **Data used in this work : 2021 L-band, 236 deg2 , ~26 scans**



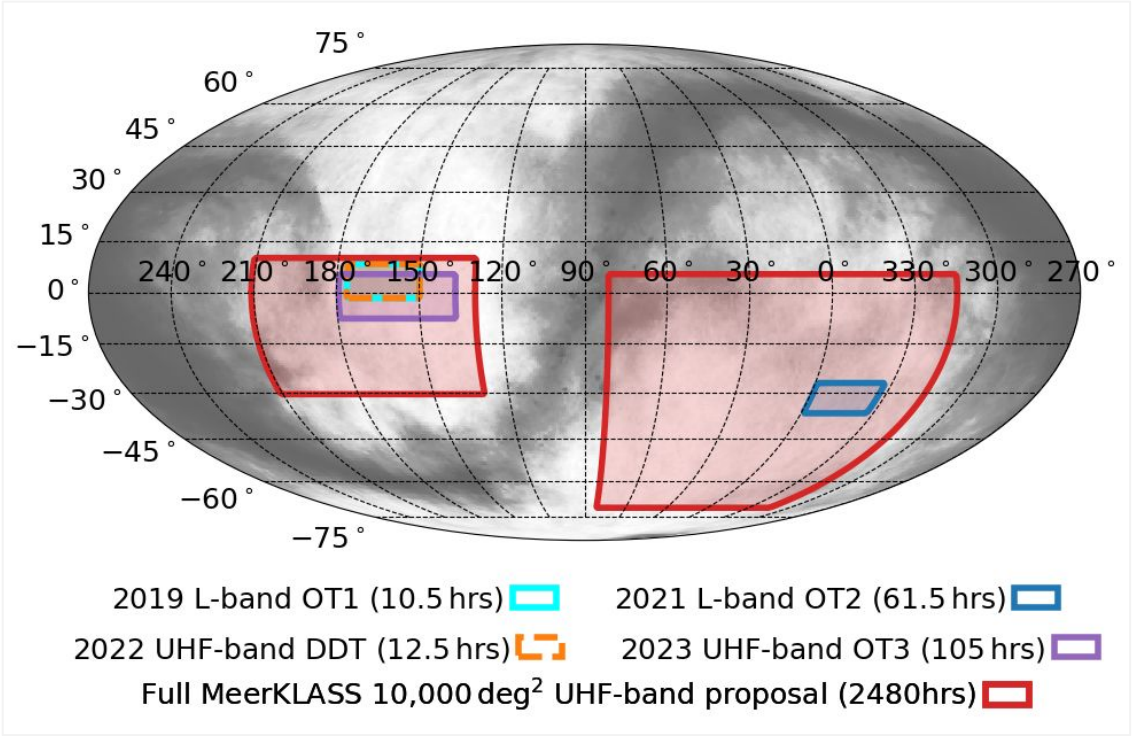


Figure 1. MeerKLASS survey footprint. Ref[meerklass.org]

The MeerKAT Radio Telescope and Its 'known' RF Environment

- Radio Frequency Interference are unwanted modulated disturbances that contaminates samples of visibility data.
- RF from satellites, GPS, Aircraft transponders and telecommunication towers (GSM) or natural occurrences of RF from electrical storms affect the Array.

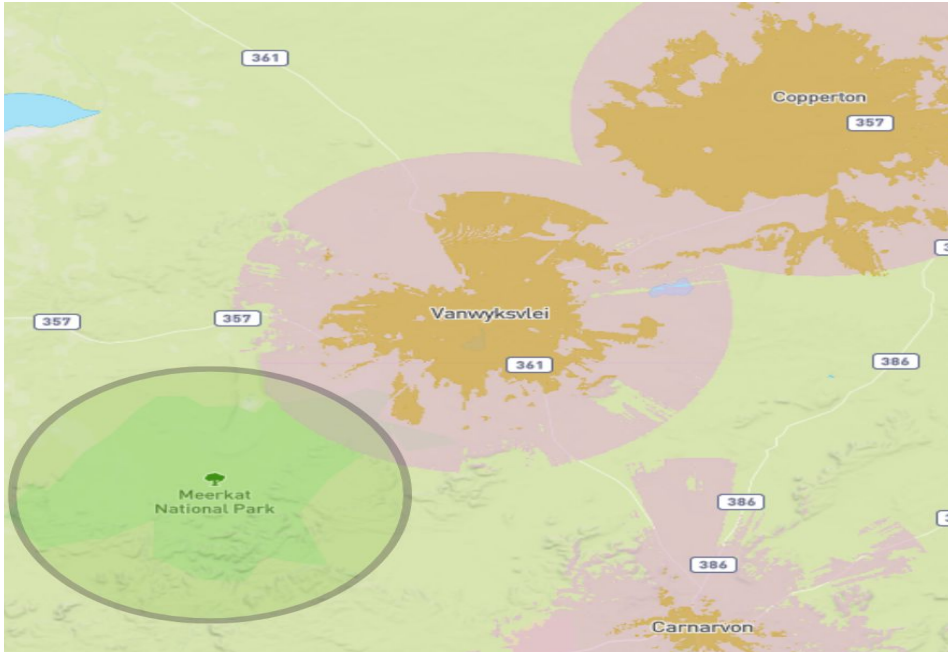


Figure 2. Cell tower network coverage around the MeerKAT

- MeerKLASS L-Band severely contaminated by RFI
- MeerKLASS KATcali pipeline:
  Calibration and data analysis pipeline Has multi-round RFI flagging of MeerKLASS datasets. Flags strong RFI known RFI in TOD to low level RFI in map domain.

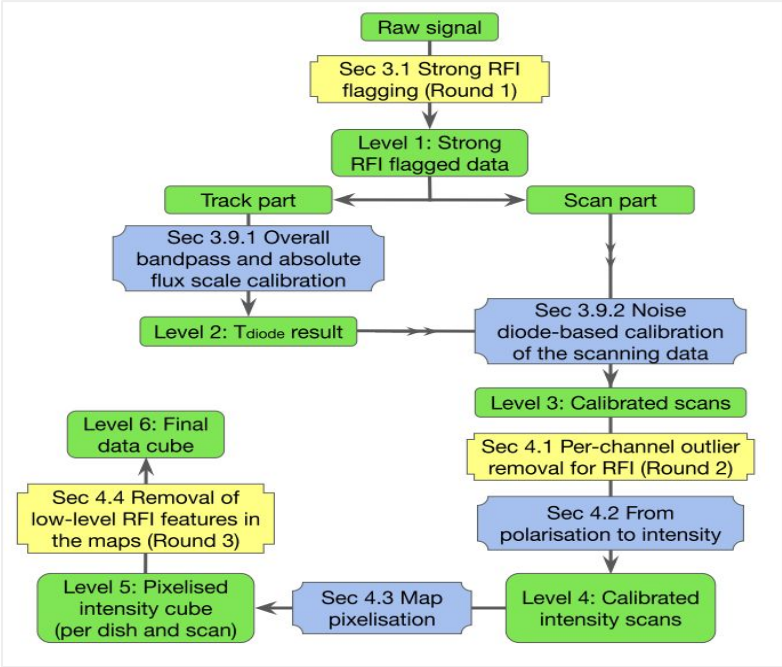
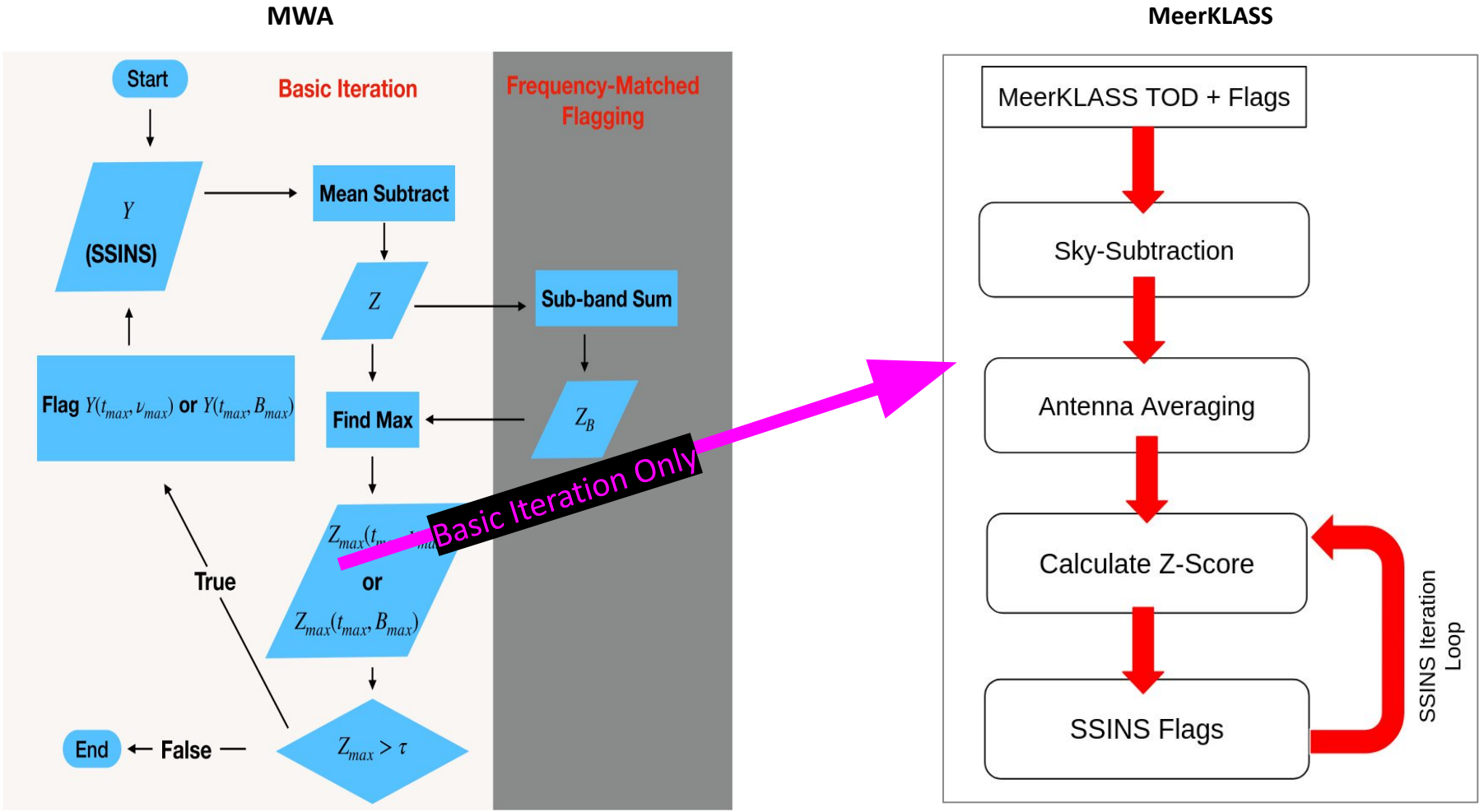
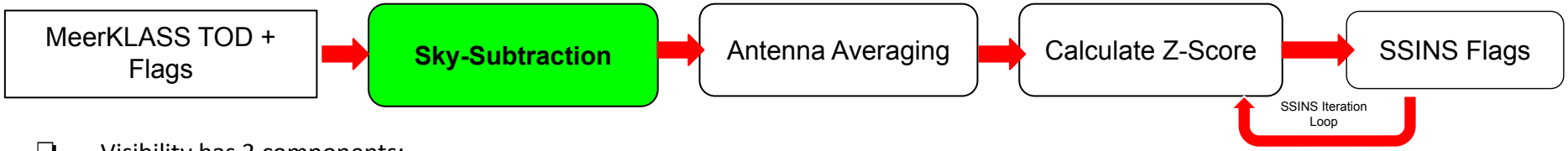


Figure 3. Flowchart of KATcali pipeline.

Sky **S**ubtracted **IN**coherent averaged **S**pectra **(SSINS)**



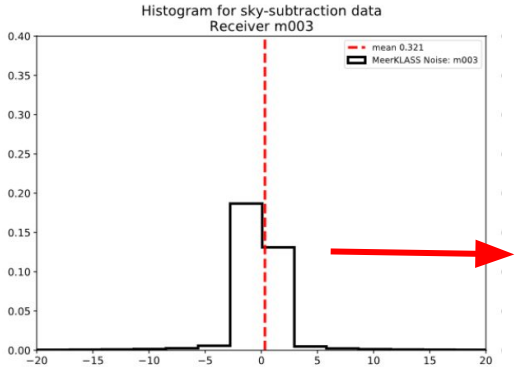


- Visibility has 3 components:
  Sky + Thermal Noise + RFI

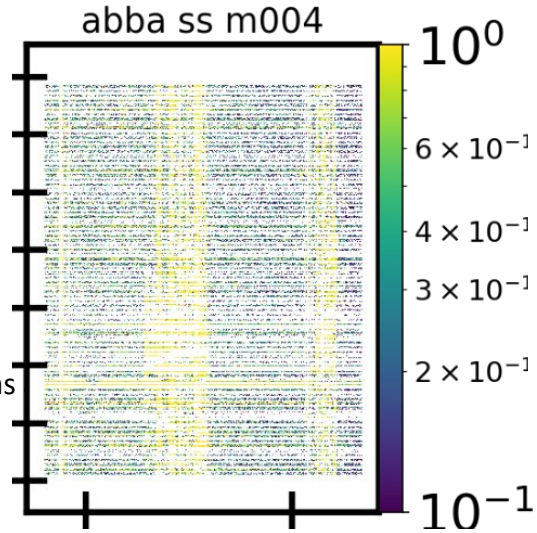
$$\Delta V_{ii}(t,\nu) = \tfrac{1}{2}\Big(V_{ii}(t_n,\nu) + V_{ii}(t_{n+1},\nu)\Big) - \tfrac{1}{2}\Big(V_{ii}(t_{n-1},\nu) + V_{ii}(t_{n+2},\nu)\Big)$$

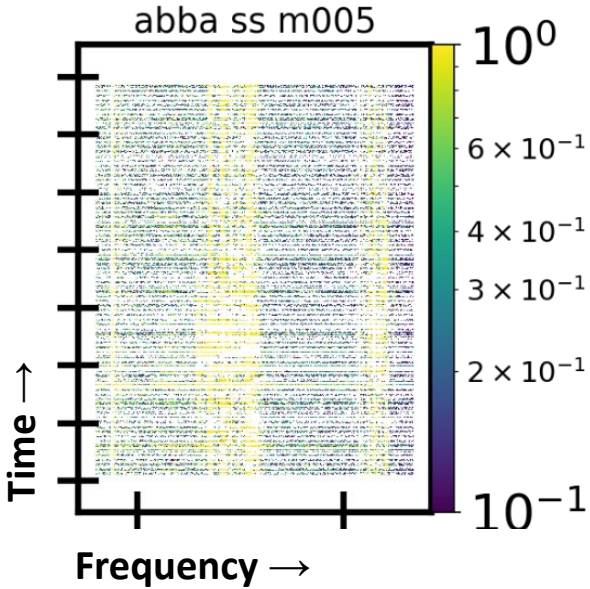
- 2s time differencing undersampled the sky subtraction for MeerKLASS data. We use ABBA difference statistic is applied to remove slowly+constantly varying sky

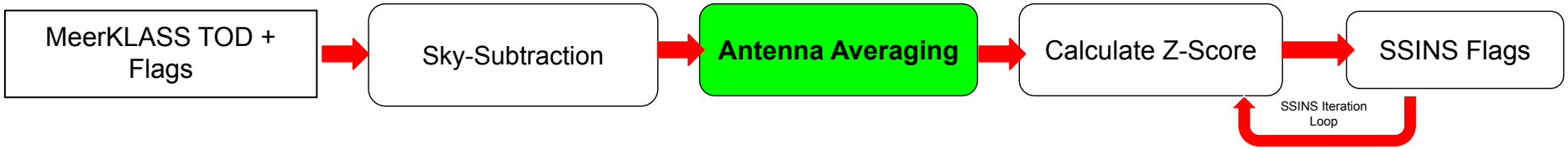
- Each receiver will have a noise spectrum, that has ~gaussian distribution.



Noise spectrum of each receiver has n approximate gaussian distribution with long tails (RFI + noise mixed distribution)







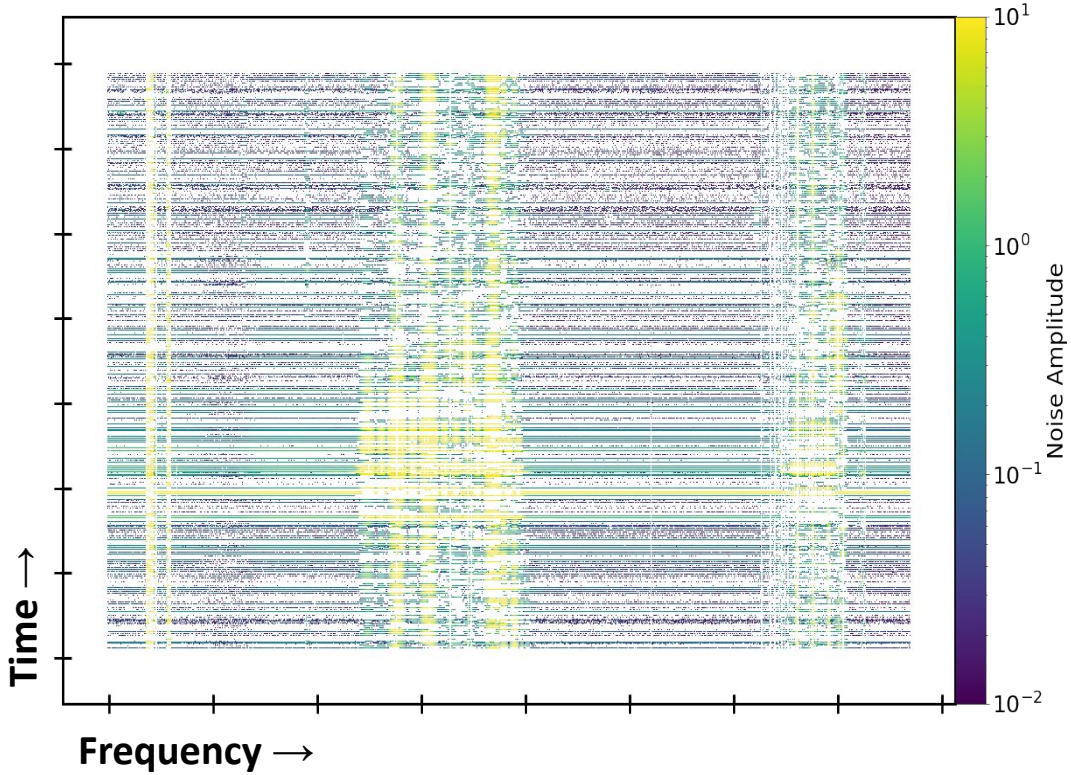
- **Average amplitude of sky-subtracted noise spectrum over all good receivers** in a single observation block

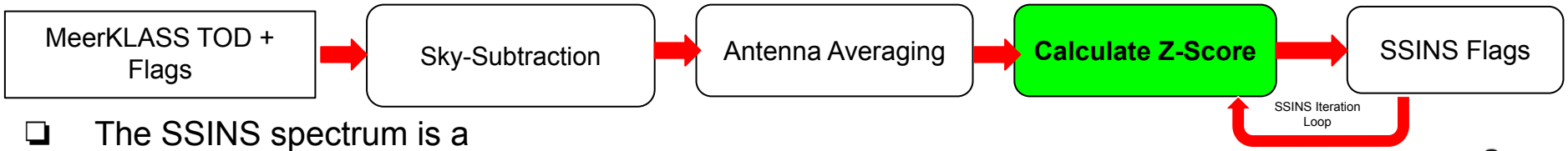
$$Y(t,\nu,p) = \frac{1}{D}\sum_{i}^{D} \Delta V_{ii}(t,\nu,p)$$

- D is the number of receivers in observation block, specifically the antennas that have level 1 checked masks avail, this is different from the list of good antennas that katcali output

- With sufficient number of receivers, the SSINS will be Gaussian distributed (derivable by the central limit theorem)

- Mixture Distribution - statistics are gaussian and an RFI distribution

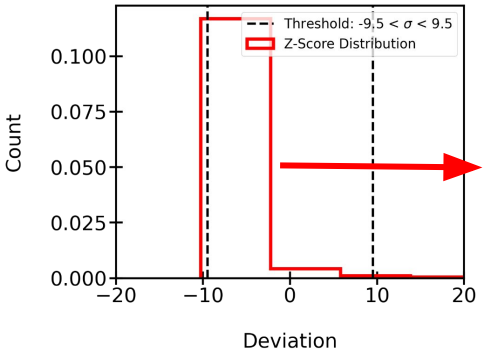




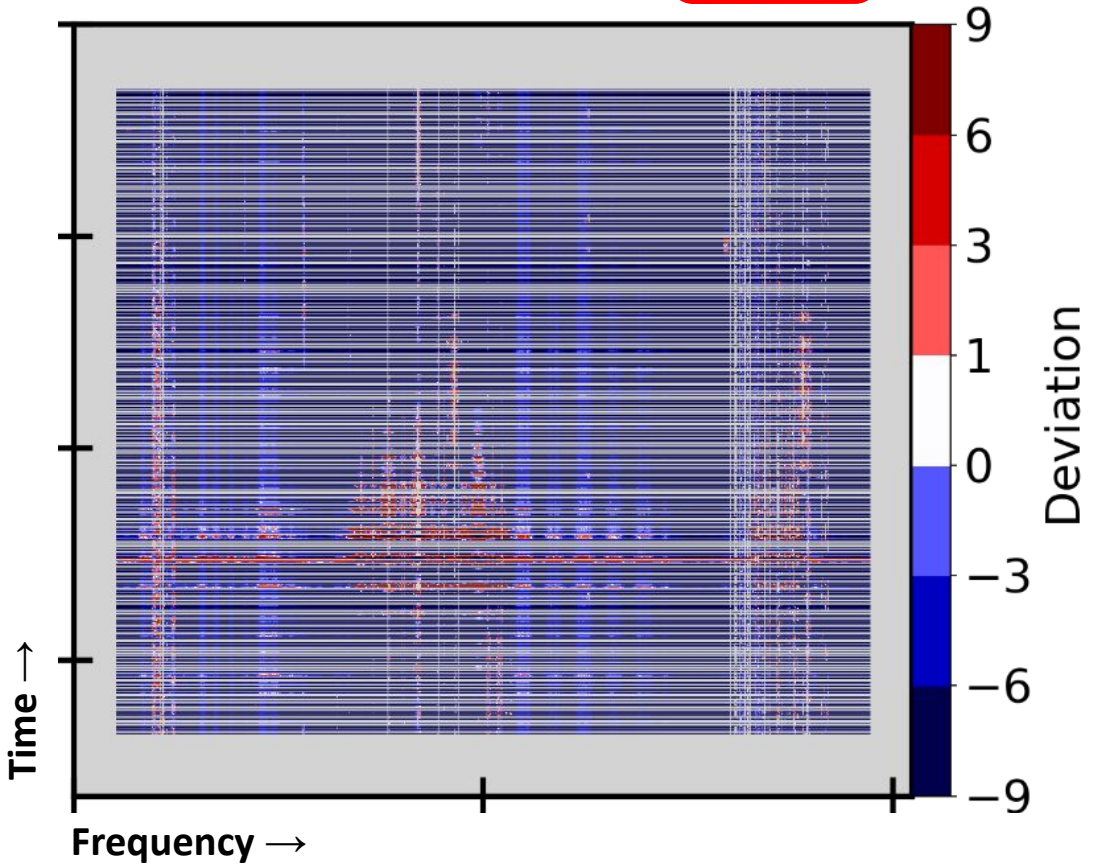
- The SSINS spectrum is a normal dist.

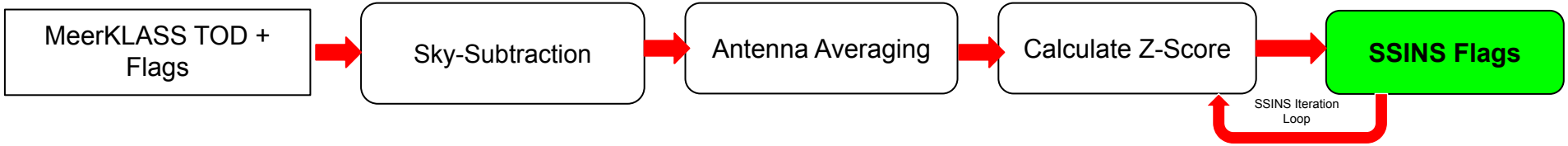
$$Z(t,\nu) = \frac{Y(t,\nu) - \mu(\nu)}{\sqrt{\frac{C \cdot \mu(\nu)^2}{D}}}$$

- Z-score measures the deviations of data points from the mean value of the distribution.
- **High +/- z scores indicate data exceeding unlikely to be thermal noise.**

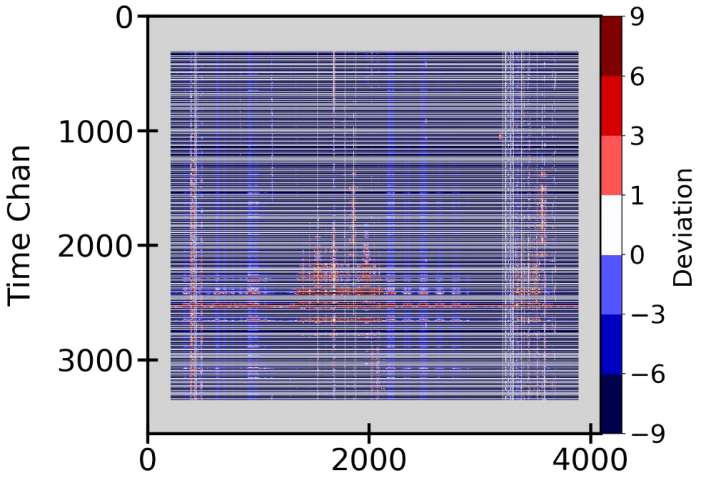


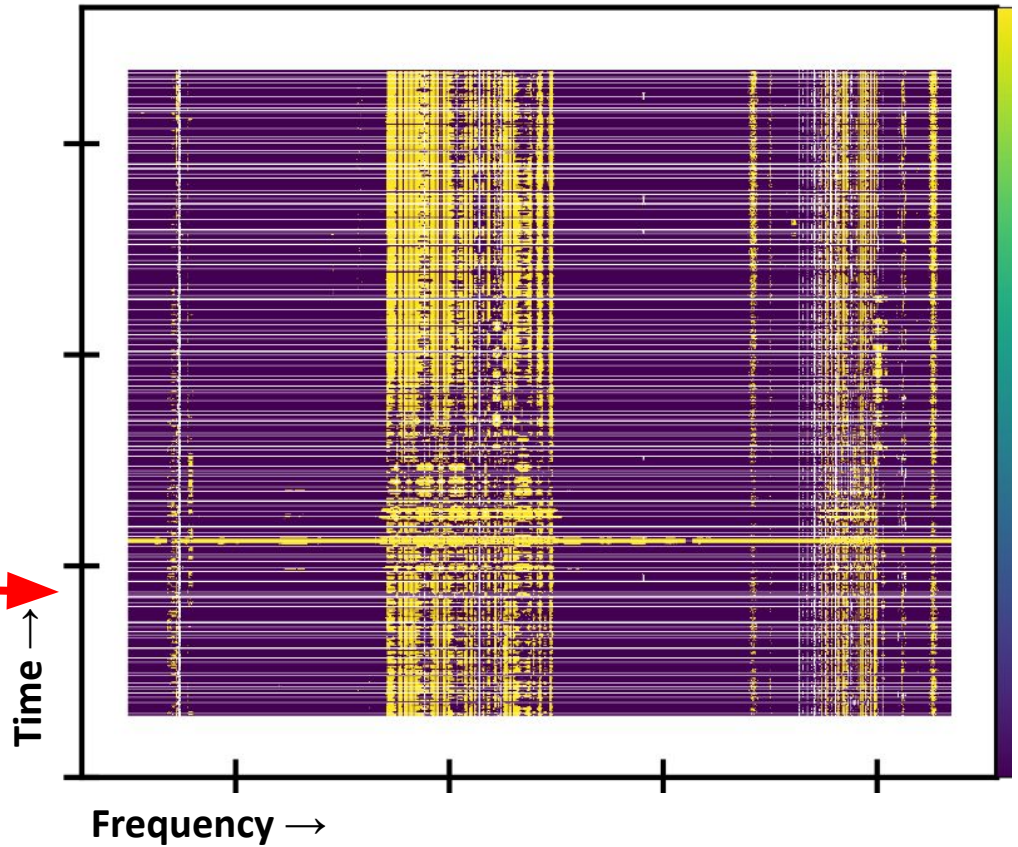
Contrasts the mixture distribution to highlight outlier data points (RFI)

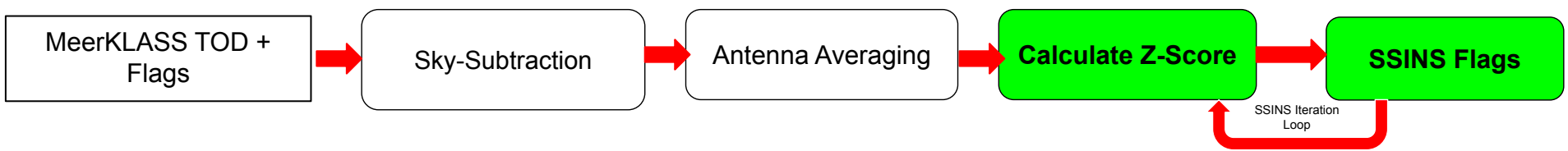




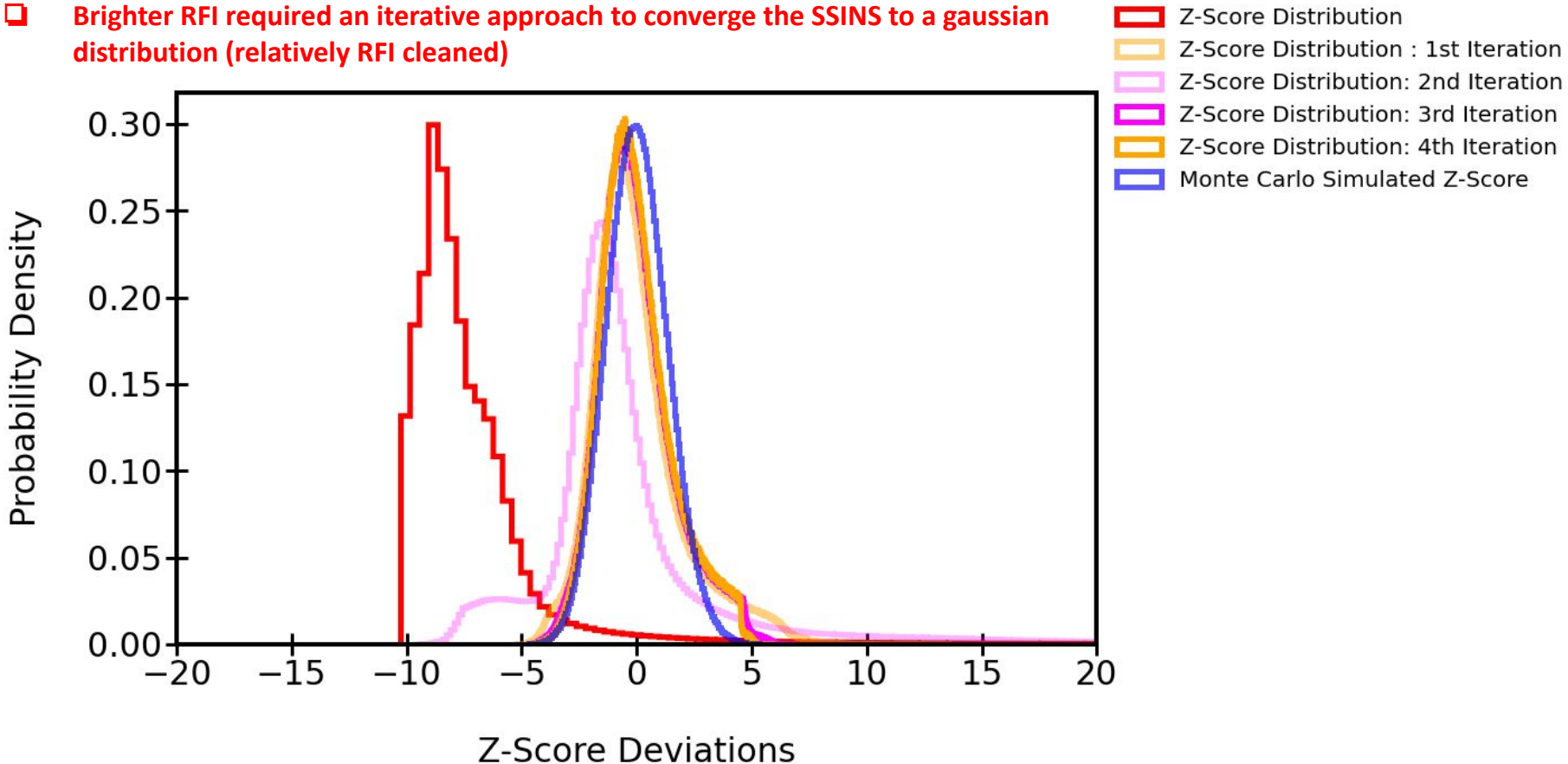
- Apply standard deviation thresholding on the z-score → get z-score masks
- We propagate z masks to TOD equivalent data flags.
- SSINS Round 1 flag occupancy is ~19%



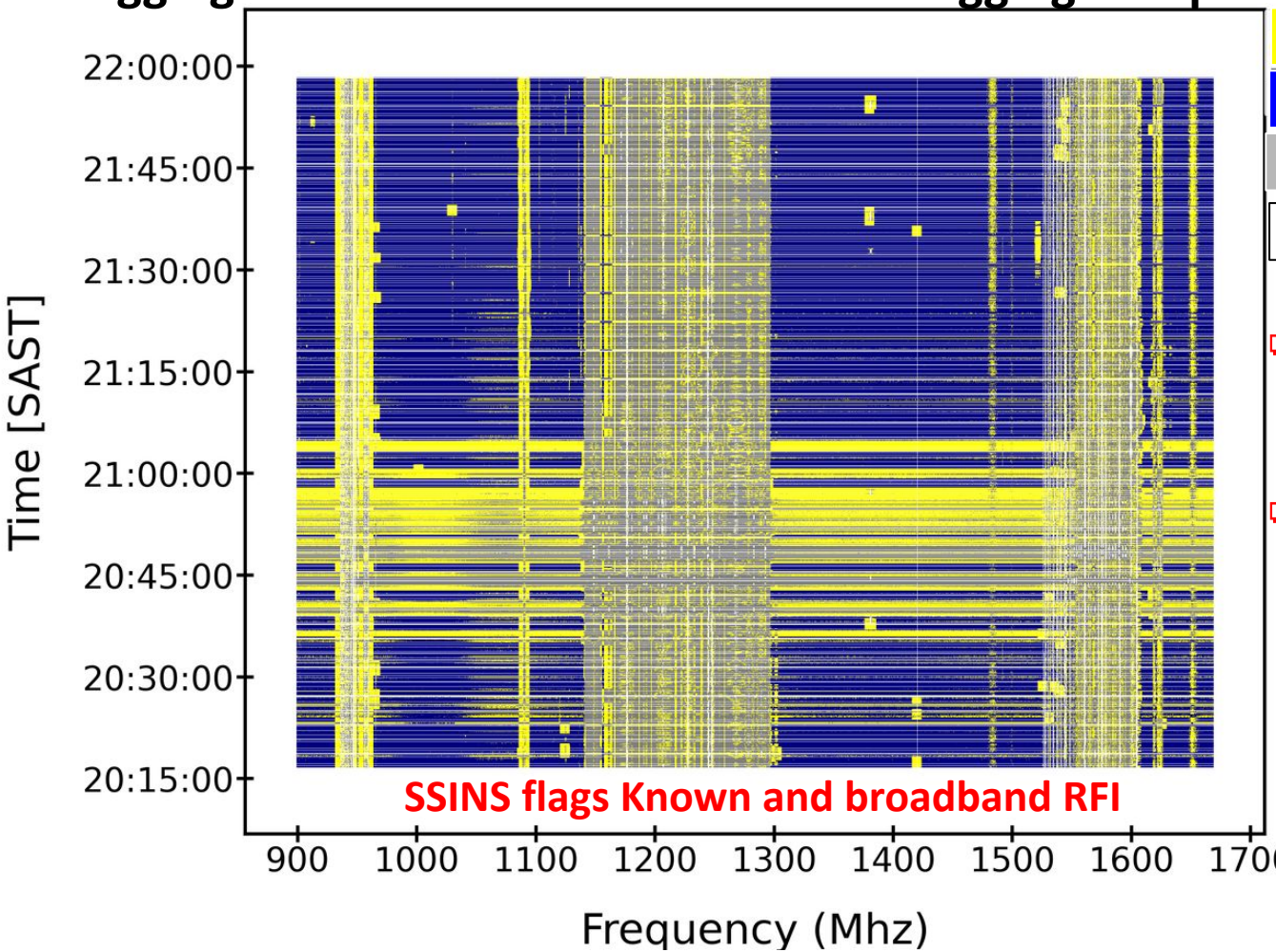




- **Brighter RFI required an iterative approach to converge the SSINS to a gaussian distribution (relatively RFI cleaned)**



Flagging Statistics : SSINS Round 1 RFI Flagging Comparison



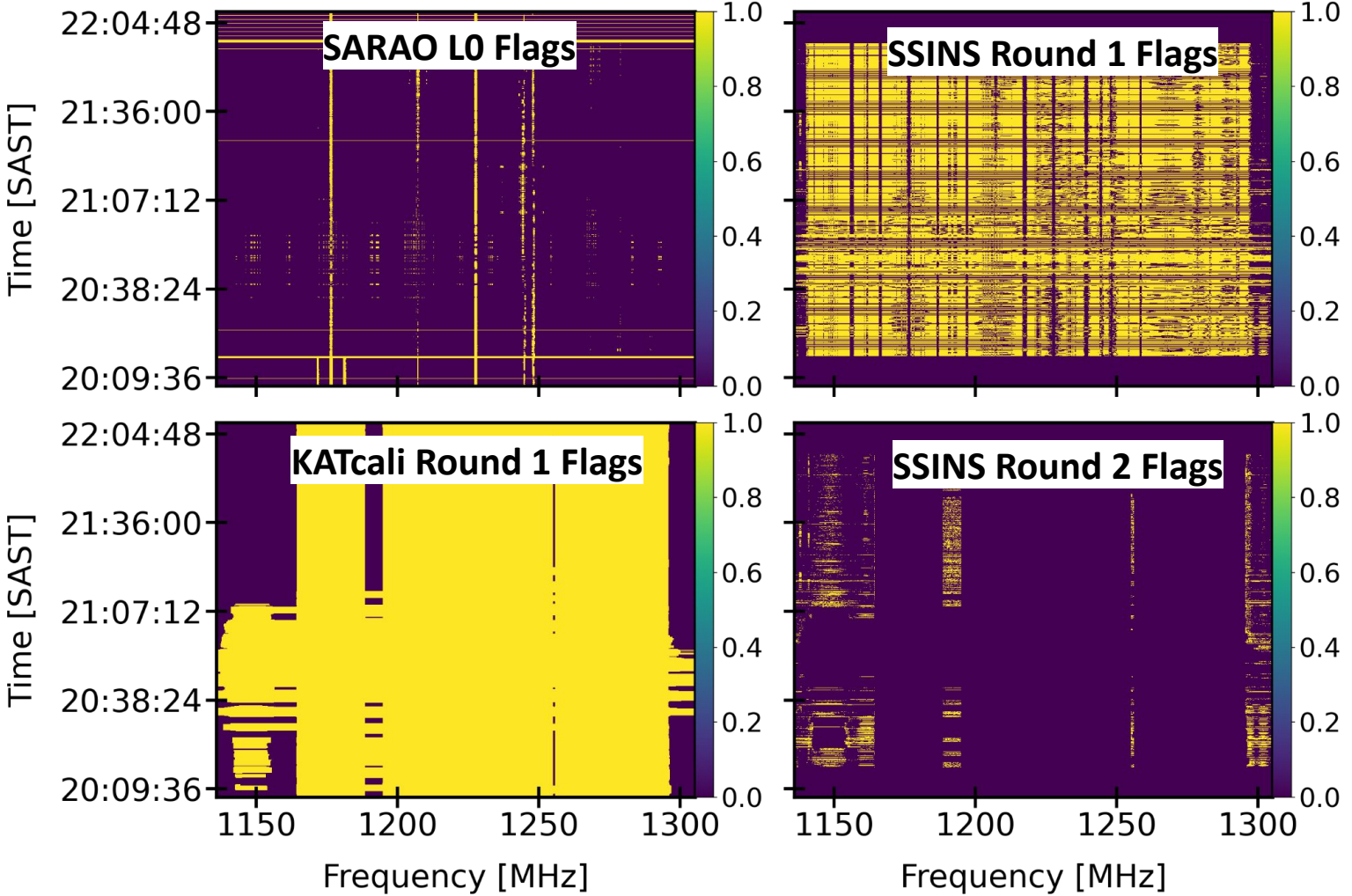
SSINS Flags : ~19.86%

Unflagged

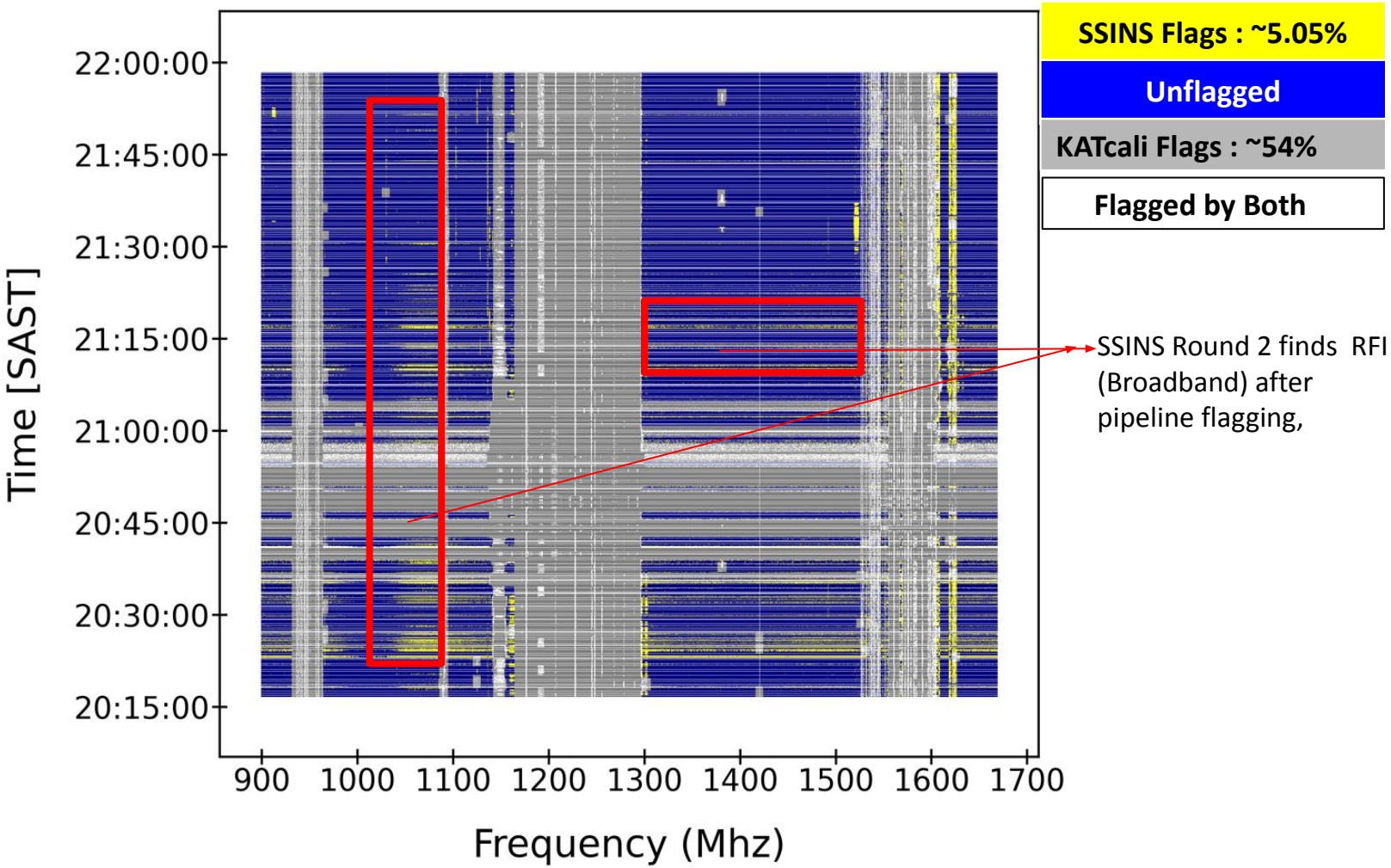
KATcali Flags : ~54%

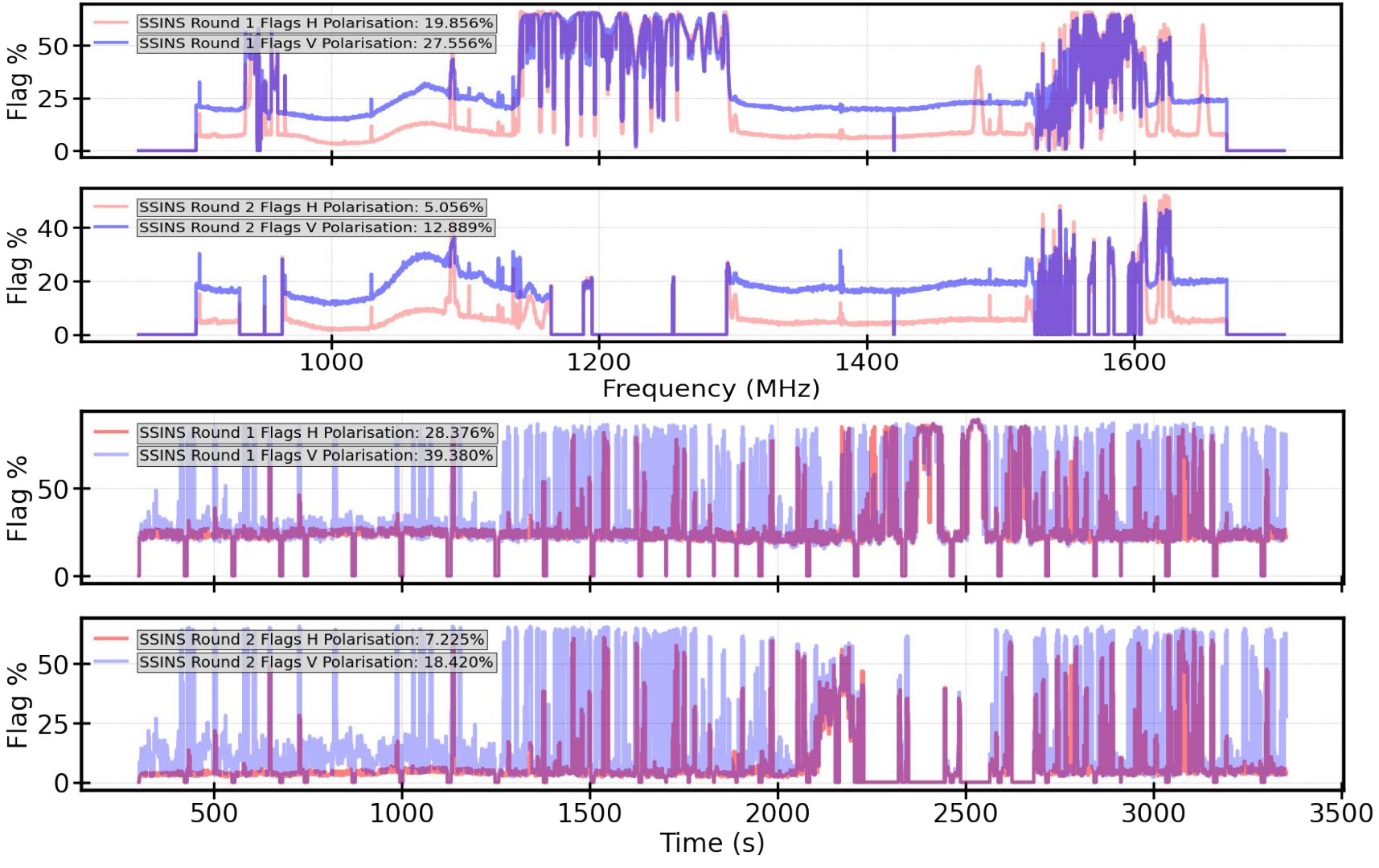
Flagged by Both

- SSINS conservative flagger- attributed to the progressive thresholding we apply.
- SSINS flags known strong RFI + extended fainter broadband RFI.


SARAO L0 Flags
SSINS Round 1 Flags
KATcali Round 1 Flags
SSINS Round 2 Flags
Time [SAST]
22:04:48
21:36:00
21:07:12
20:38:24
20:09:36
1150
1200
1250
1300
Frequency [MHz]
0.0
0.2
0.4
0.6
0.8
1.0

Flagging Statistics : SSINS Round 2 RFI Flagging Comparison

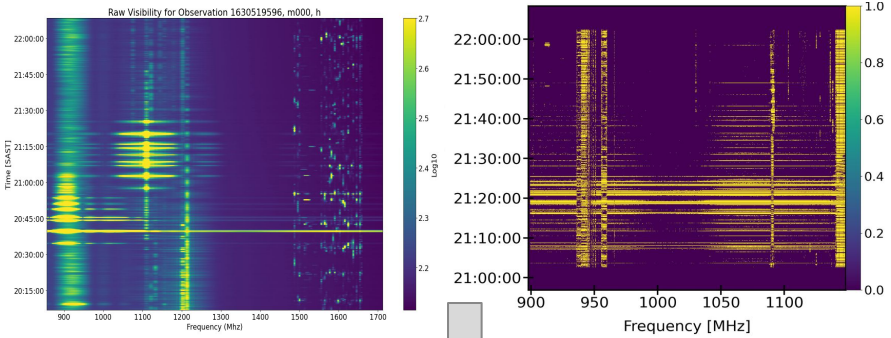


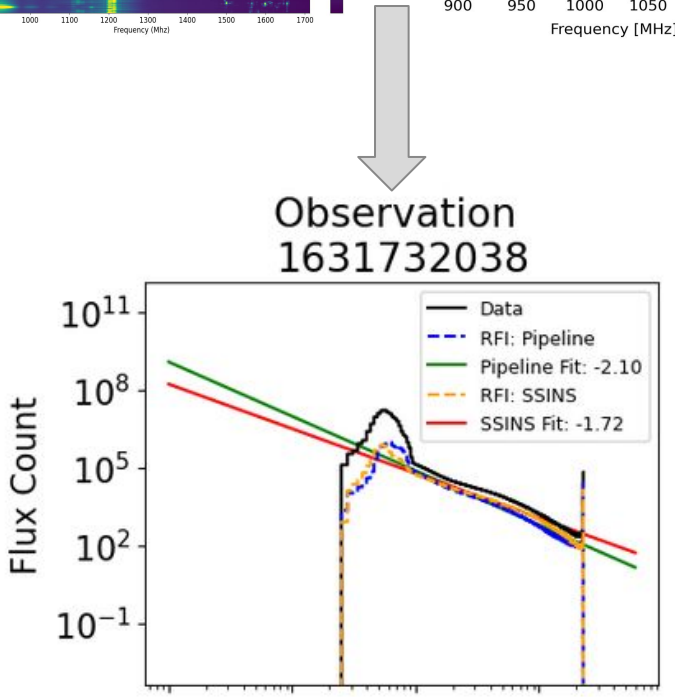

SSINS Round 1 Flags H Polarisation: 19.856%
SSINS Round 1 Flags V Polarisation: 27.556%
Flag %
SSINS Round 2 Flags H Polarisation: 5.056%
SSINS Round 2 Flags V Polarisation: 12.889%
Flag %
Frequency (MHz)
SSINS Round 1 Flags H Polarisation: 28.376%
SSINS Round 1 Flags V Polarisation: 39.380%
Flag %
SSINS Round 2 Flags H Polarisation: 7.225%
SSINS Round 2 Flags V Polarisation: 18.420%
Flag %
Time (s)

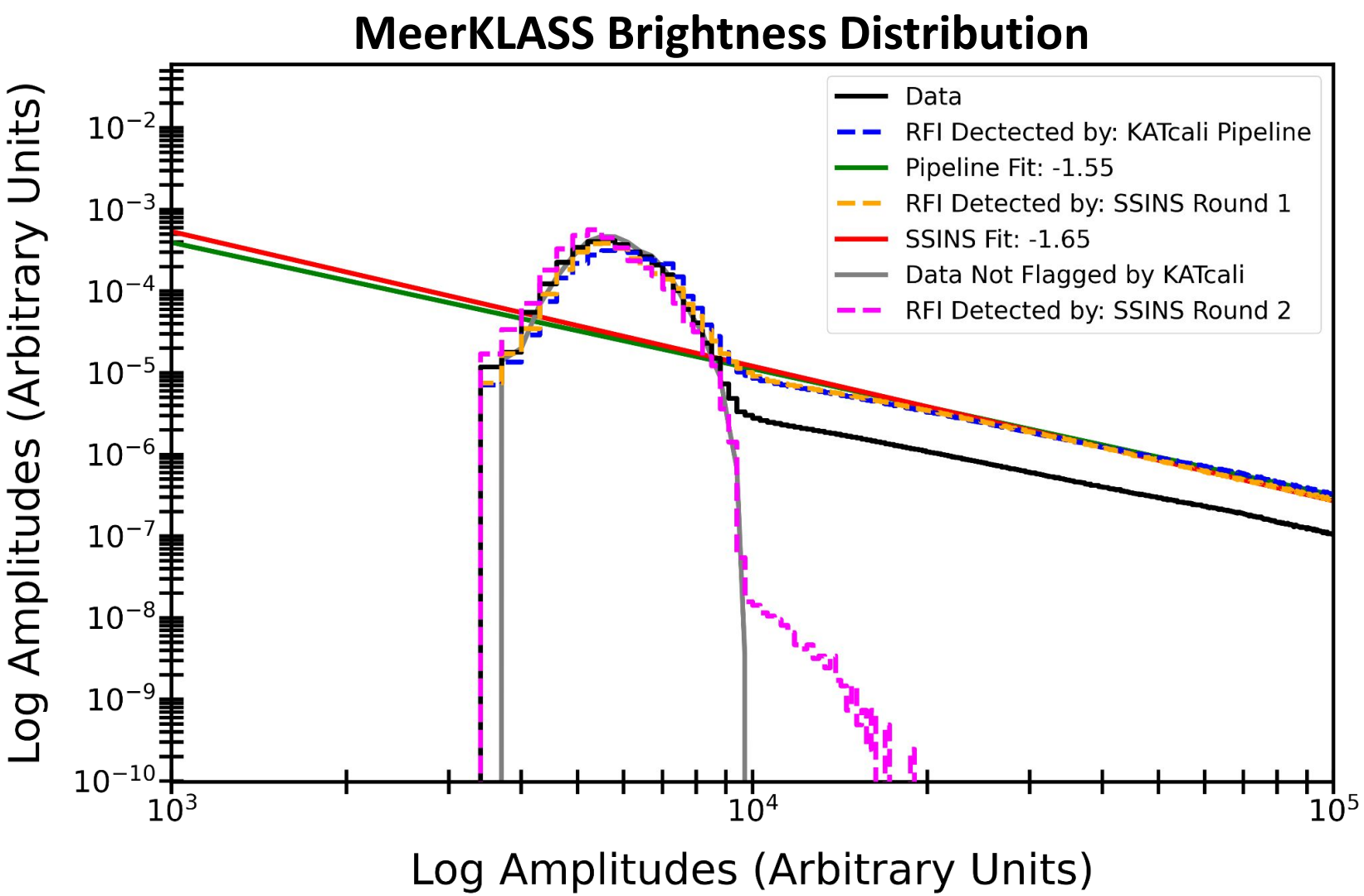
Brightness Distribution Analysis

- Based on radio source counts (LogN - LogS Analyse) done in cosmology studies.
  - RFI emitters are treated as celestial sources
  - counts the rate density of the interference, revealing statistical footprint of the intensity the array accumulated.
  - Does undetected RFI integrate over many hours?

- We use the brightness distribution as a statistical tool to:
  - Analyse the unflagged/ flagged data, we expect that thermal noise (follows a gaussian) and RFI (empirical power-law, highly non-thermal behaviour)
  - Investigate how well we have flagged RFI.






MeerKLASS Brightness Distribution
Data
RFI Dectected by: KATcali Pipeline
Pipeline Fit: -1.55
RFI Detected by: SSINS Round 1
SSINS Fit: -1.65
Data Not Flagged by KATcali
RFI Detected by: SSINS Round 2
Log Amplitudes (Arbitrary Units)
Log Amplitudes (Arbitrary Units)
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-10}
10^{3}
10^{4}
10^{5}

Summary

- Results suggest that SSINS is quite an effective RFI flagger. In most case, it can replace R1 katcali flags while providing additional low-level flags.

- Comparative assessment between katcali and SSINS flags and their histogram distributions implies that we are doing pretty good job on RFI flagging
- Future Outlooks
  - Implement SSINS on UHF data
  - Investigate how the brightness distribution can inform RFI characterizing

meerklass.org