RFI Mitigation for the ngVLA – a cost-benefit analysis

Urvashi Rau, Rob Selina, Alan Erickson

National Radio Astronomy Observatory

IEEE APS-URSI 2020

RFI mitigation for the ngVLA

(1) RFI Landscape

(2) Mitigation approaches integrated with system design

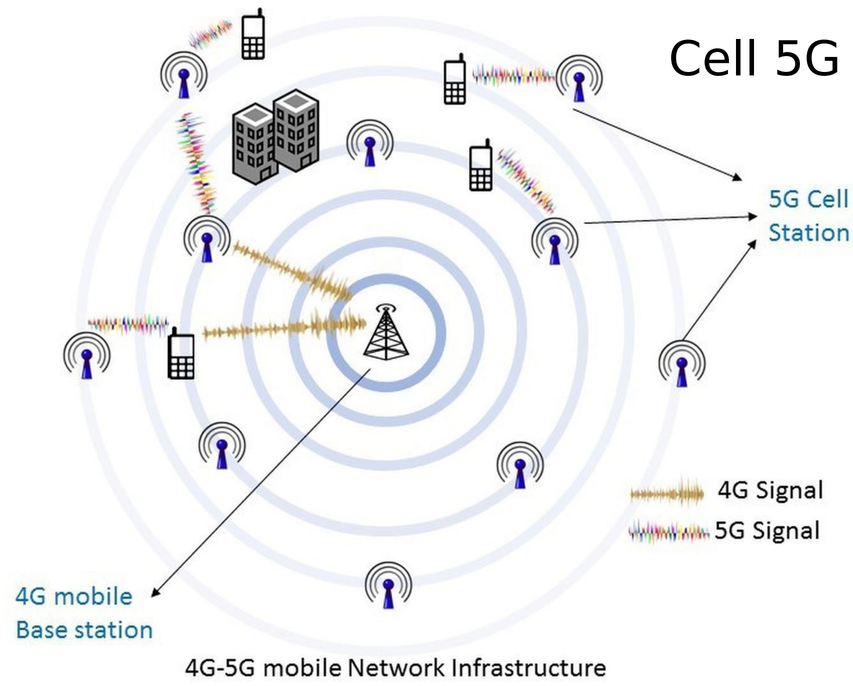
(3) A cost benefit analysis to identify areas of focus

Details : ngVLA Memos #48, #70, #71

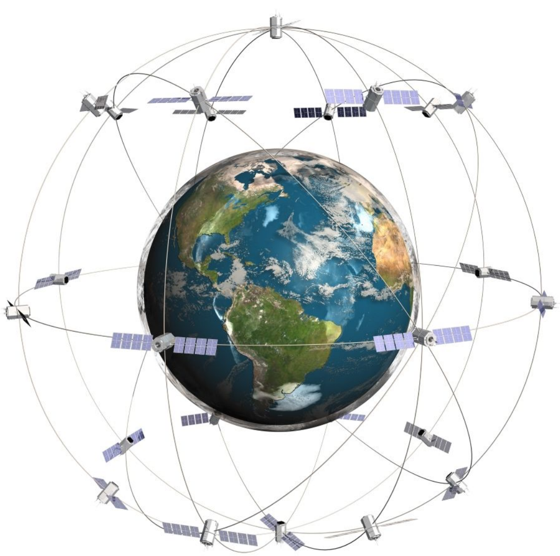
Radio Frequency Interference (RFI)

Commercial signals transmitted in the same frequency range as interesting astrophysical emissions.

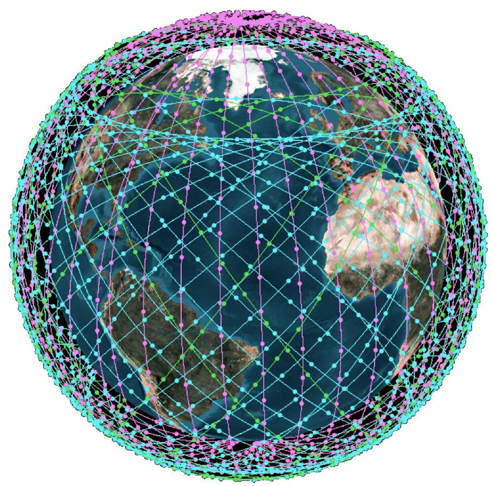
Cell 5G



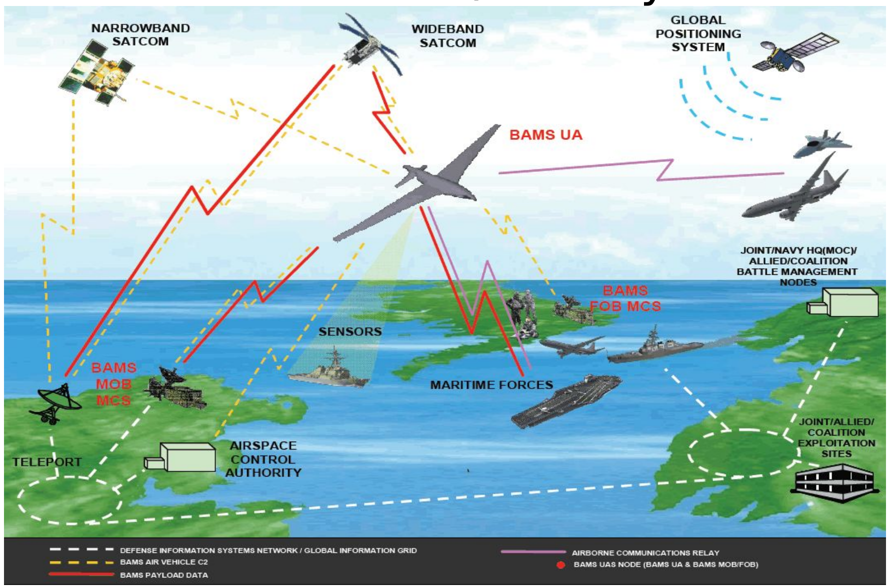
GPS-sat



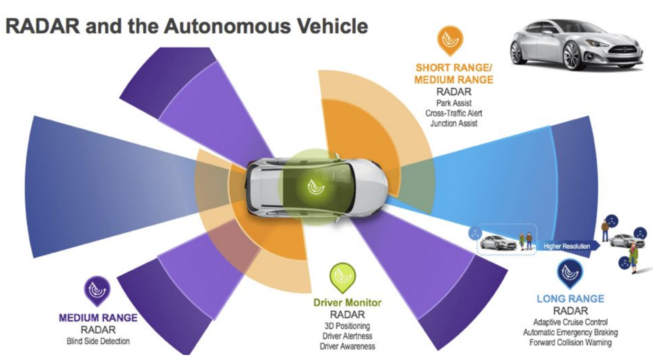
LEO-sat



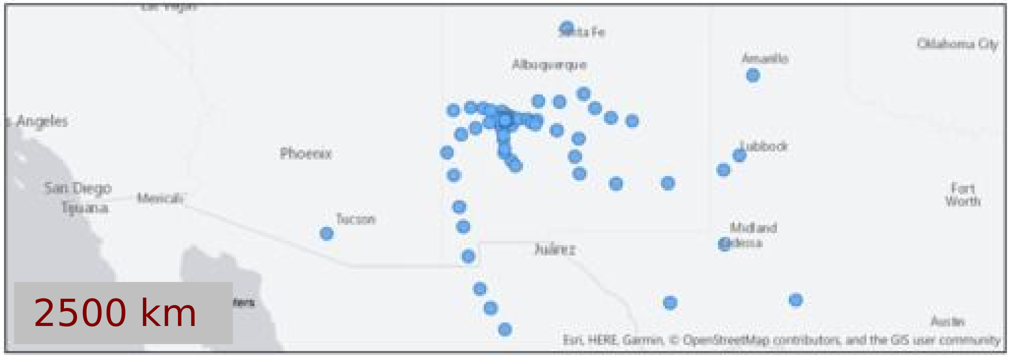
Aviation/Military Comm

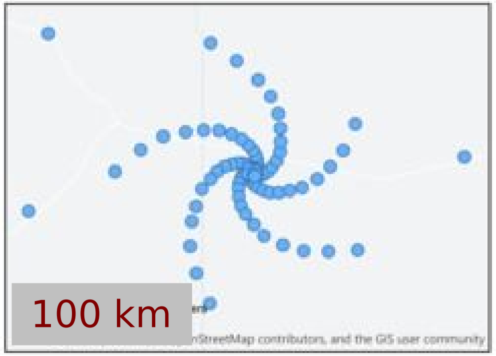


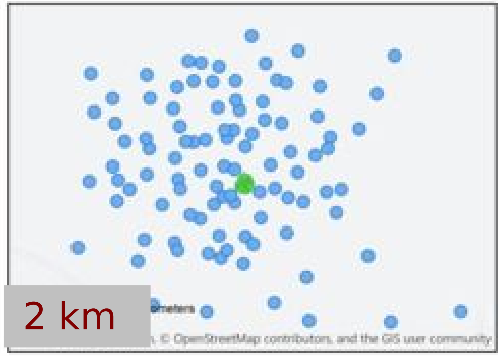
Car Radar

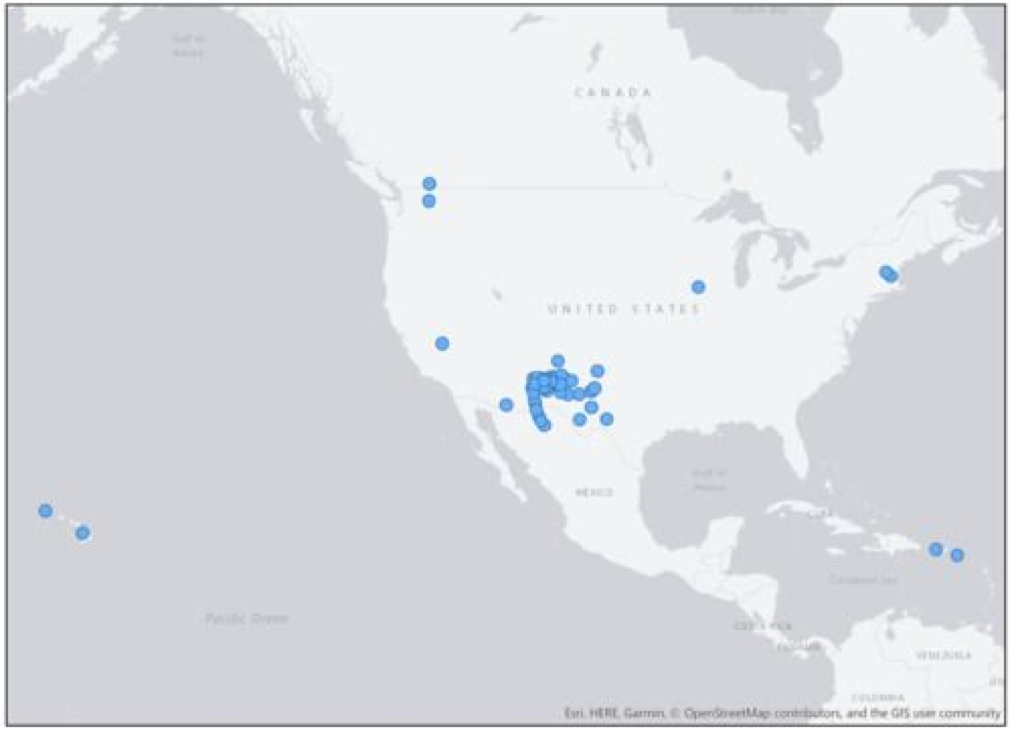


Next Generation VLA









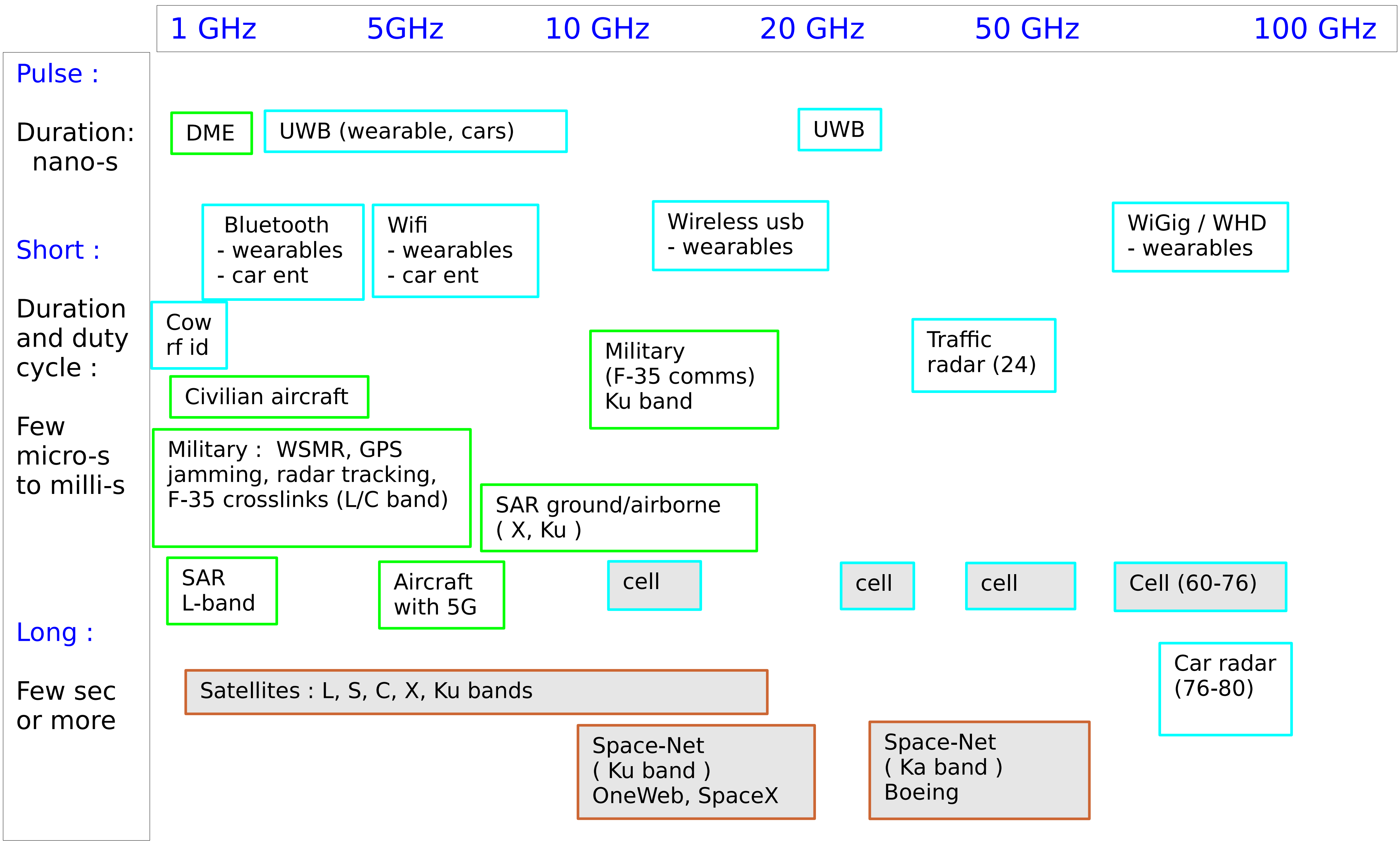
Frequency range : 1.2 – 116 GHz

Main Core : 100 x 18m dishes (< 2km baselines)
Plains Spiral : 114 x 18m dishes (< 70 km baselines)
Compact core : 19 x 6m dishes (< 0.1 km baselines)
Long Baselines : 30 x 18m dishes (100 – 8000 km baselines)

Main/Compact Cores (119) : Plains of San Augustin (remote)

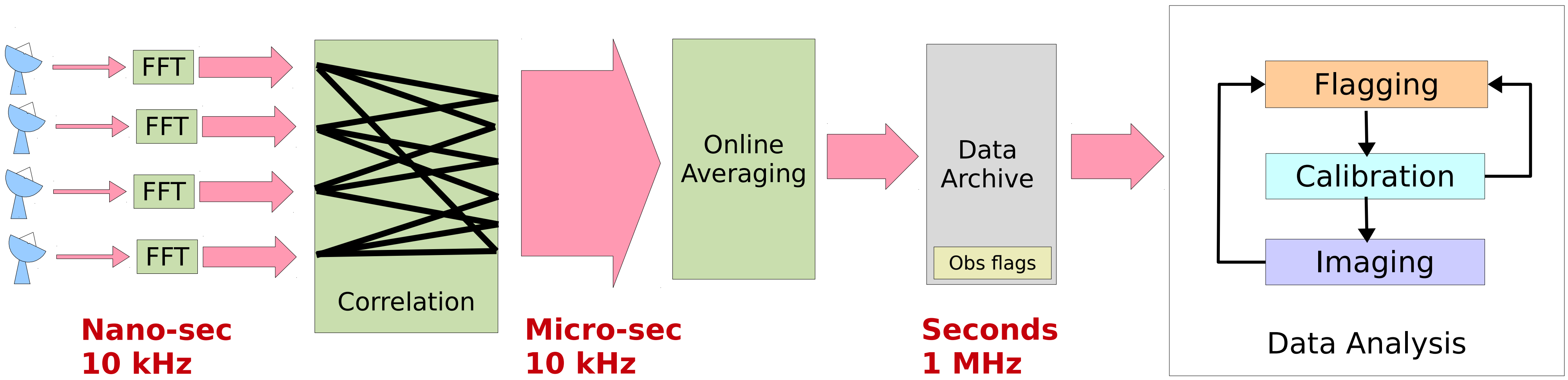
Other antennas (144) : Near humans. All antennas (263) : Will see satellites

Future RFI Landscape (1-100 GHz)

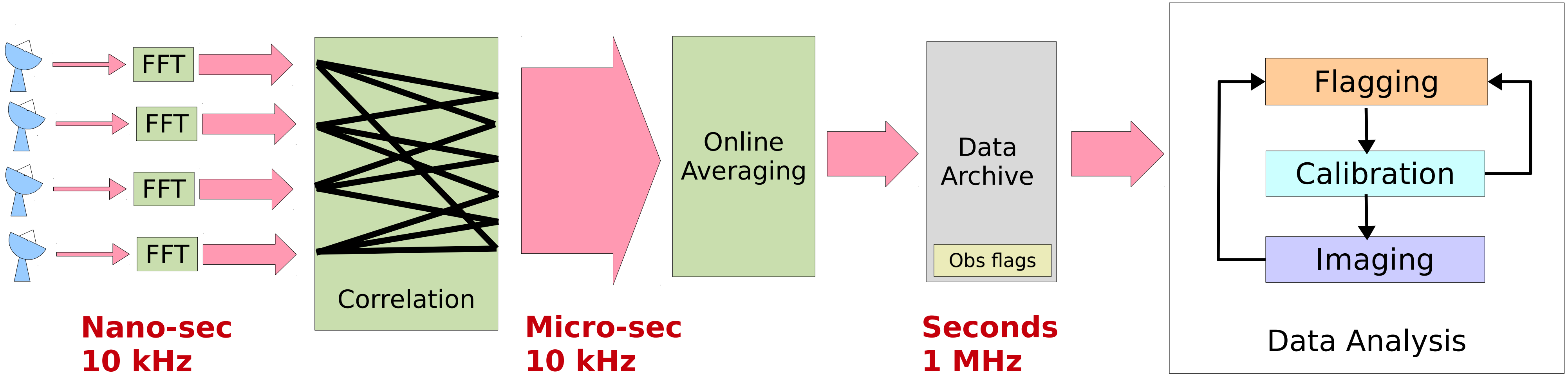


Color : Local RFI (~ few antennas) RFI on large fraction of array(airborne) RFI on entire array (satellite)
Shading : White : Seen for a small fraction of observing time. Grey : Seen for most/all observations

Data acquisition and analysis



Data acquisition and analysis



Problems :

- At 1sec, 1 MHz resolutions, intermittent RFI appears continuous

  => We are throwing away good data

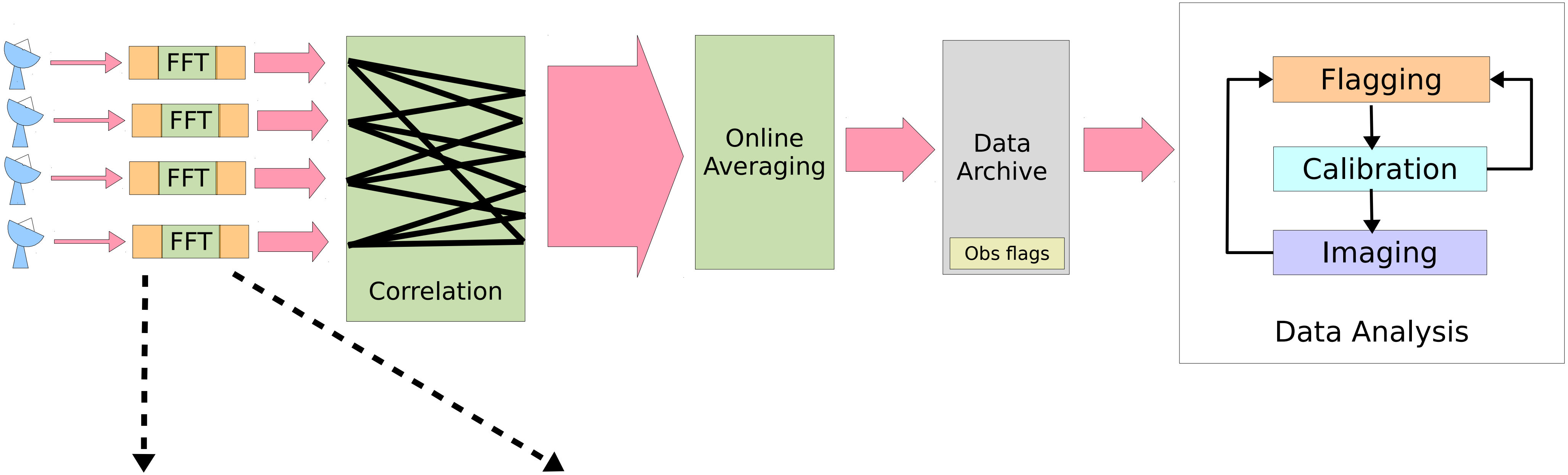
- Satellite RFI is partially decorrelated at 1sec, 1 MHz resolutions

  => Cannot model the signals well enough to subtract them

- No interaction between scheduling and post-processing or science needs

  => Not making use of what we learn.

Real-time RFI mitigation before correlation



Outlier detection on time-series.

Detect and mask nano-second pulses seen by individual antennas

RFI : Sparks,
UWB transmissions

Action : Drop Samples
Gather metadata

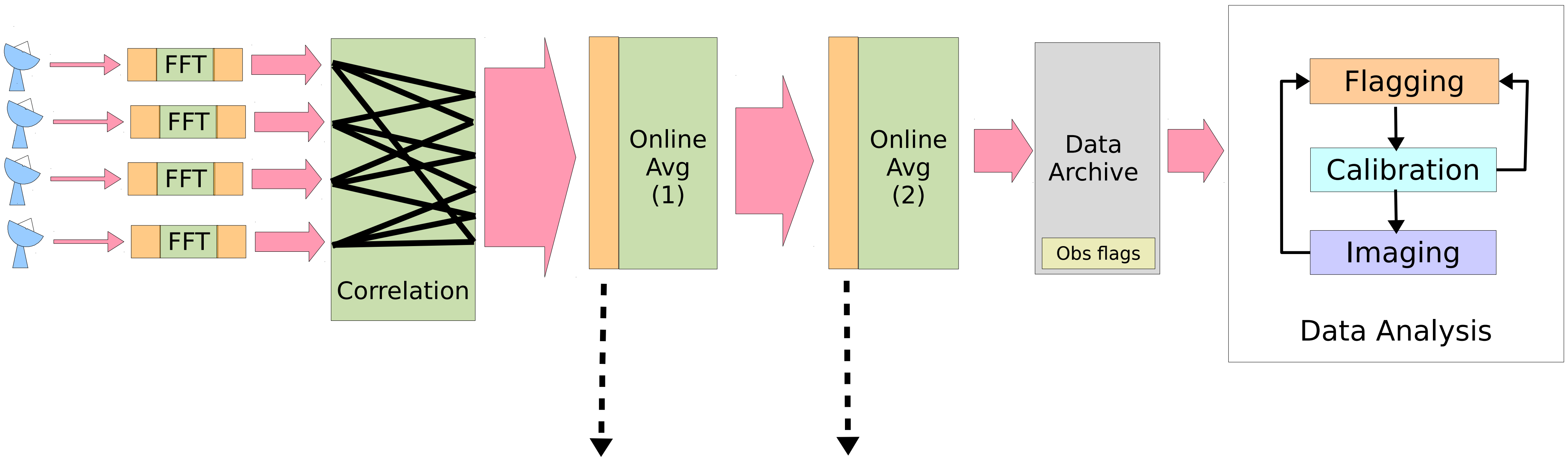
Outlier detection on real-time spectrum

Detect and mask/label corrupted channels per antenna

RFI : Any strong narrow-band signal

Action : Mask samples
Label samples
Gather metadata

Real-time RFI mitigation after correlation



Automatic flagging on visibilities
(1 micro-sec, 10 KHz)

Detect and mask correlated RFI in the baseline x time x frequency (3D) cube

RFI : Intermittent communication signals

- Duty cycles of 10s of micro-sec
- Channel width of 100 KHz.

Action : Drop samples during averaging
Record correct weights

Interference modeling and subtraction
(1 milli-sec, 10 kHz : no/partial decorrelation)

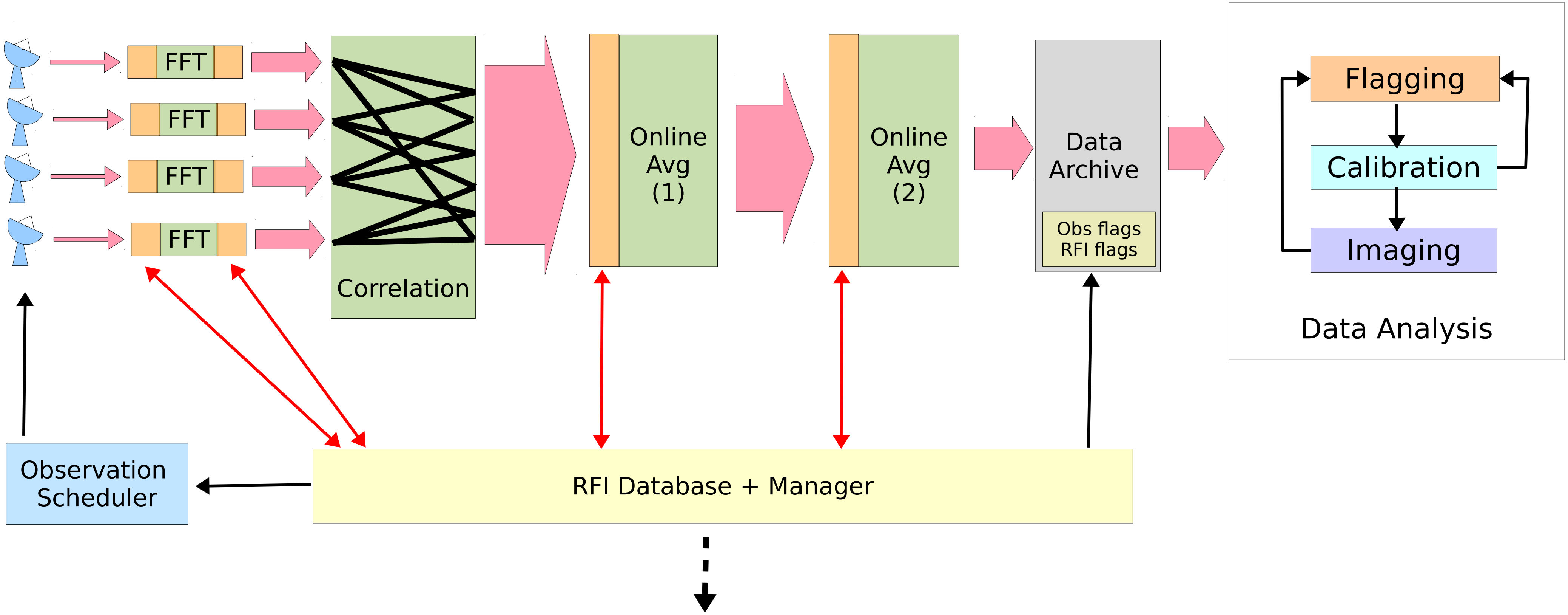
Matrix subspace projection.
Real-time all-sky Imaging. Source location.

RFI : Continuous data transfer signals

- Cell 5G, Airborne 5G
- Satellite Radio / Internet

Action : Subtract out the RFI signal
Modify weights

RFI Database and Manager



Database : Store RFI characteristics and meta-data

- Known satellite orbits and frequencies, locations / schedules of terrestrial emitters,
- Meta-data about RFI detected by the real-time system

Manager : Analyse RFI metadata and decide optimal actions for the current observation.

- Record RFI information in the archive
- Match the current RFI to suitable mitigation algorithms and tunings
- Smart scheduling around predictable (or currently detected) emitters

Moving from theory to the real world…..

A high level architecture diagram.

A variety of ideas

– Evaluate the cost versus the benefit of each solution

– Identify focus areas for proof-of-concept prototyping

=> Start with a Fermi estimation

Cost-Benefit Analysis

Factors :

- RFI characteristics (time, frequency, space)
- Algorithm options and expected efficacy
- Data Processing Stages (time/freq resolution at which the data are viewed)

---- Match algorithms to the data processing stage.

Metrics :

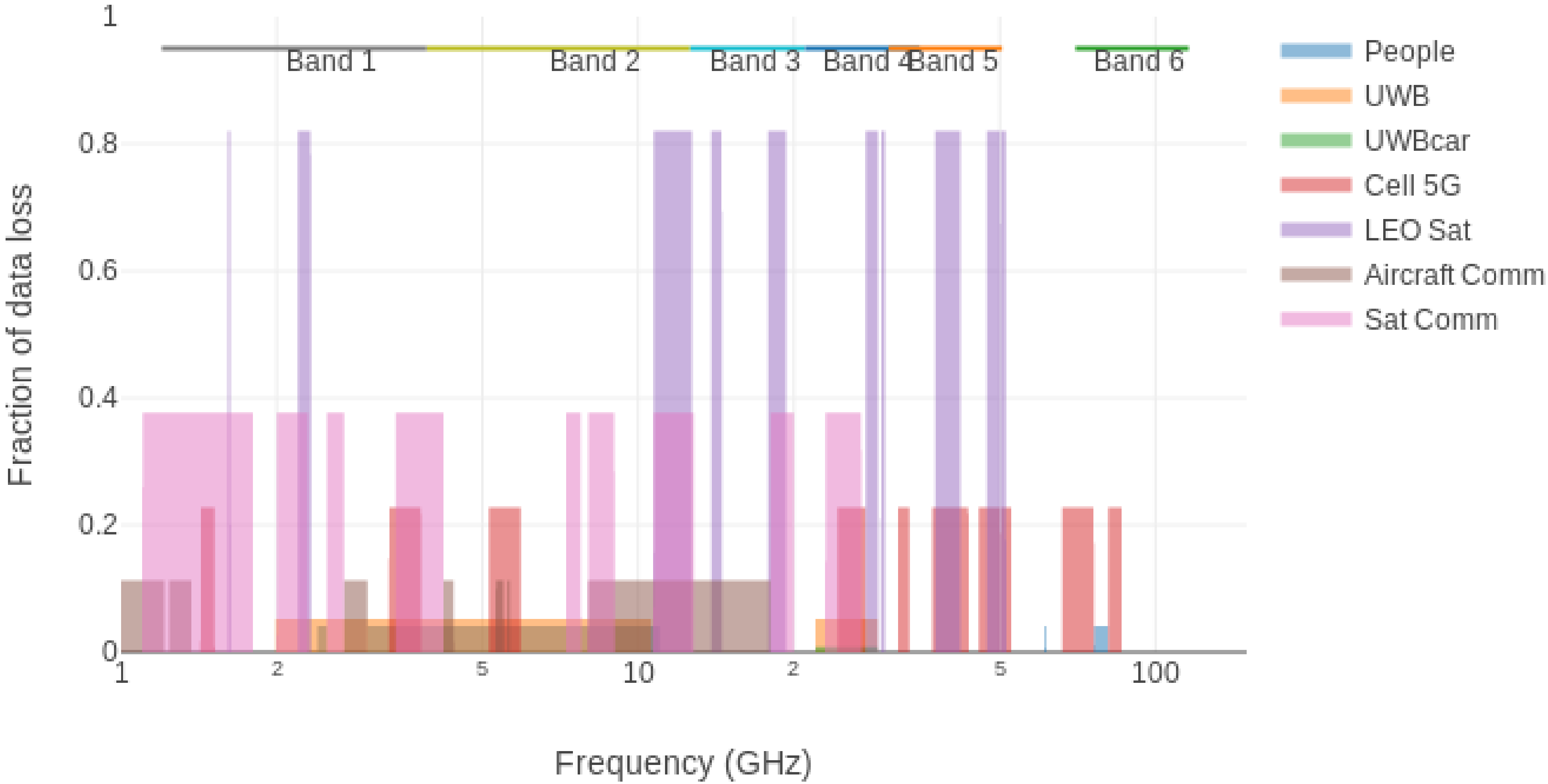
- Fraction of data loss, for different mitigation options and data resolutions.
- Extra observing time required to compensate (when possible).
- Compute cost of running chosen algorithms at the optimal data resolution

Questions :

- Is the cost of implementing and running a mitigation option going to be worth it in terms of saved observing time or science ?

- How stable are these calculations when we vary input RFI characteristics ?

Estimated data loss – only post-processing

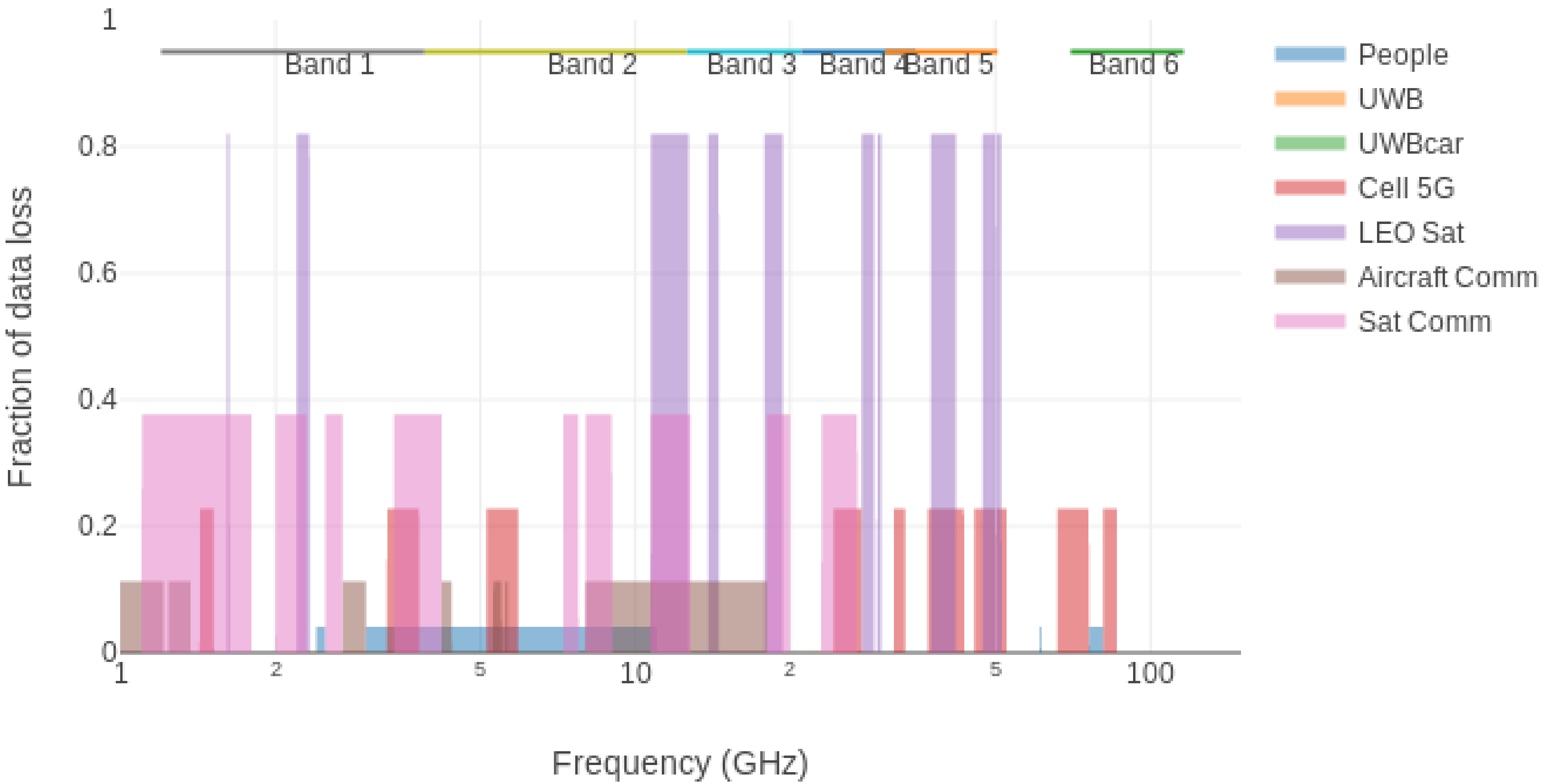


Assumptions : Multiple RFI types with different footprints in frequency, time, and antennas.
Entire allocated band is filled at once (i.e. no usable gaps), no spillover/saturation

Calculations : Fraction of affected baselines, effects of RFI decorrelation and uncorrelated RFI
RFI filling fraction (signal duty-cycle vs data resolution), overlapping vs disjoint RFI

RFI mitigation : Only post-processing flagging

Estimated data loss – RFI mitigation at the antenna

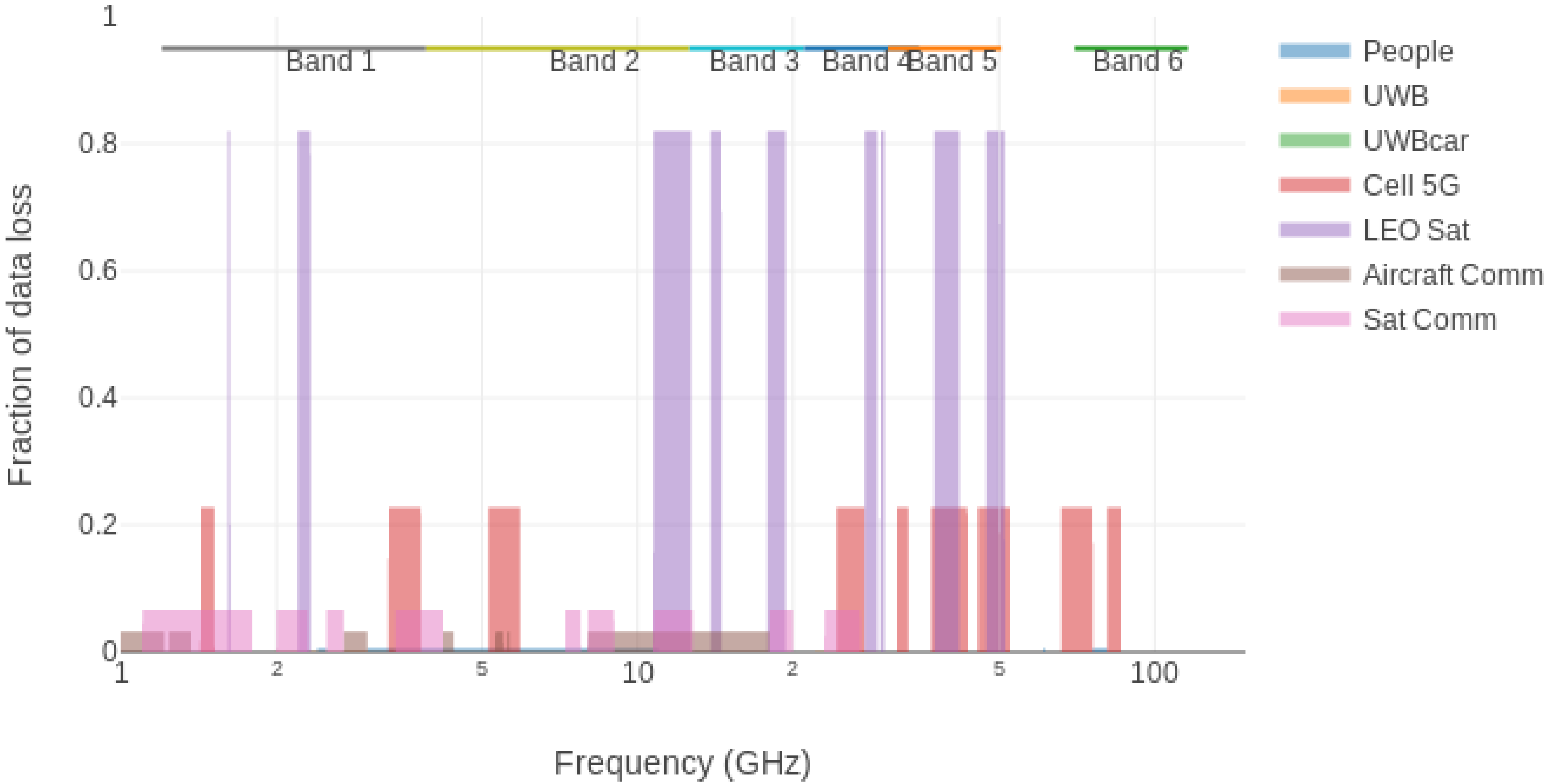


Impulsive, local broadband RFI is removed.

Low impact overall, but useful because of the UWB nature of sparky RFI.

All longer duration RFI persists

Estimated data loss – In-correlator RFI flagging

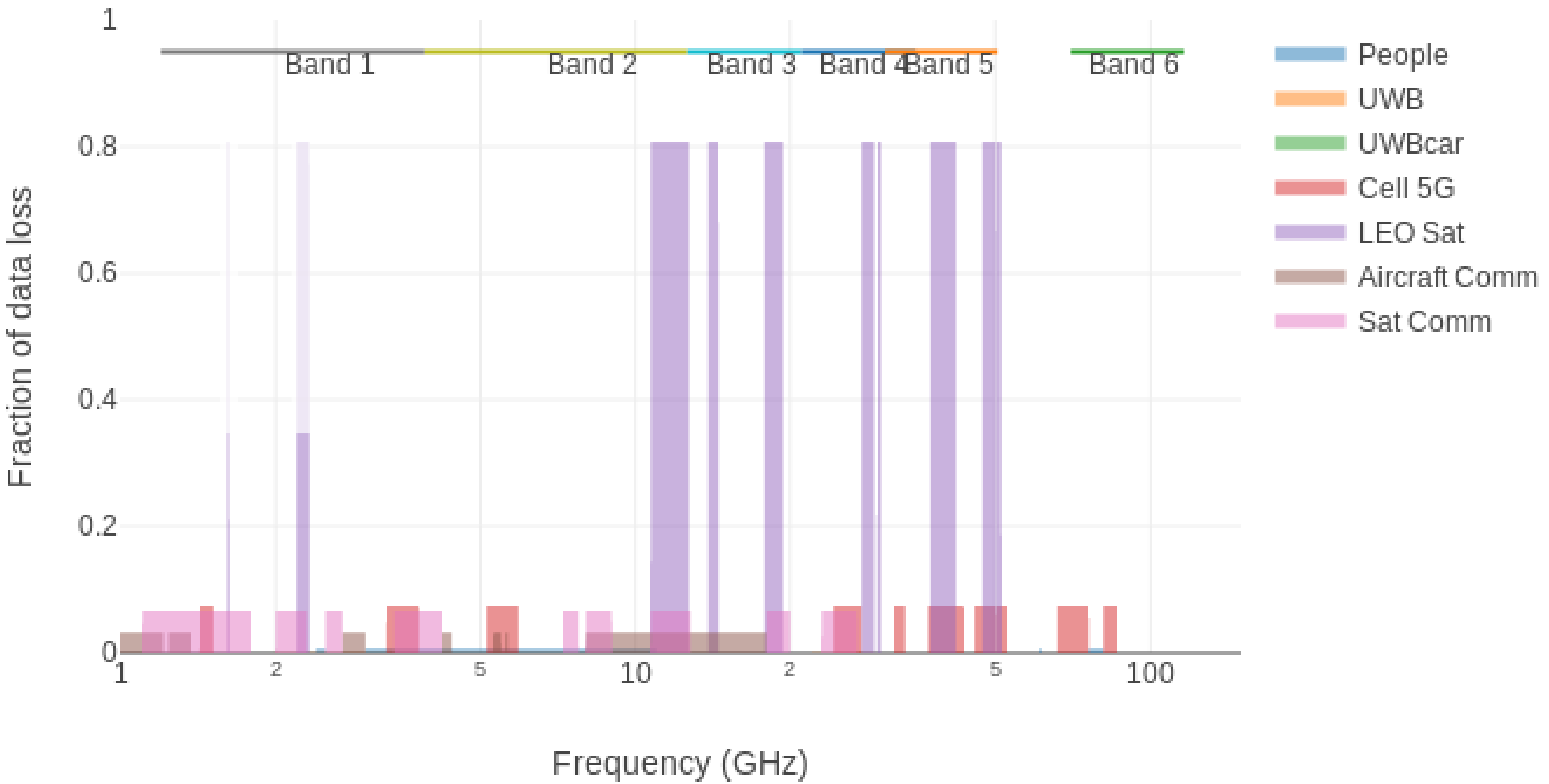


Takes advantage of duty-cycle gaps at the micro-sec to milli-sec timescale

Problem/Question : Do we really have usable gaps ?

Continuous RFI persists : Cell 5G and Satellite Data

Estimated data loss – RFI modeling and subtraction



Models and subtracts continuous RFI signals

Problem : Experiments so far have **not** been very successful

Geostationary/Local RFI is easier. LEO satellite bands may be lost (too many).

How can regulation help ?

(1) LEO satellites : An avoidance zone (footprint) above the telescope

– Main Goal : To avoid saturating entire receiver bands.

=> Data loss is confined to LEO bands only.

(2) 5G Cell Towers :

– No new 5G towers near the ngVLA array core.

- Data loss will be similar to that from LEO satellites (at diff freqs) if cell 5G is active near the core.

– Band-selection for cell 5G towers near ngVLA-spiral antennas.

(3) Other (hard to regulate, but what most of our solutions depend on) :

– The presence of RFI gaps in time and frequency

(4) Protected Radio Astronomy bands :

– Don’t lose them

RFI Impact Simulator

RFI Impact Simulator for the Next Generation VLA

RFI Characteristics

Type of RFI	Array Visibility	Time Fraction	Sig Len (s)	Sig Gap (s)	Chan Width (kHz)	
People	['outlier', 'core']	[0.9, 0.4]	0.001	0.01	20000.0	
UWB	['outlier', 'core']	[0.9, 0.2]	1e-09	1e-07	900000.0	
UWBcar	['outlier', 'core']	[0.2, 0.03]	1e-09	1e-07	8000000.0	
Cell 5G	['outlier', 'core', 'full']	[1.0, 0.2, 0.2]	0.0001	0.0	200.0	
LEO Sat	['full']	[0.9]	0.0001	0.0	200.0	[[1.61, 1.63], [2.2, 2.33], [10.7, 12.
Aircraft Comm	['core', 'outlier']	[0.5, 0.9]	2e-05	5e-05	100.0	[[1.24, 1.37
Sat Comm	['full']	[0.4]	2e-05	0.0001	100.0	

RFI mitigation options

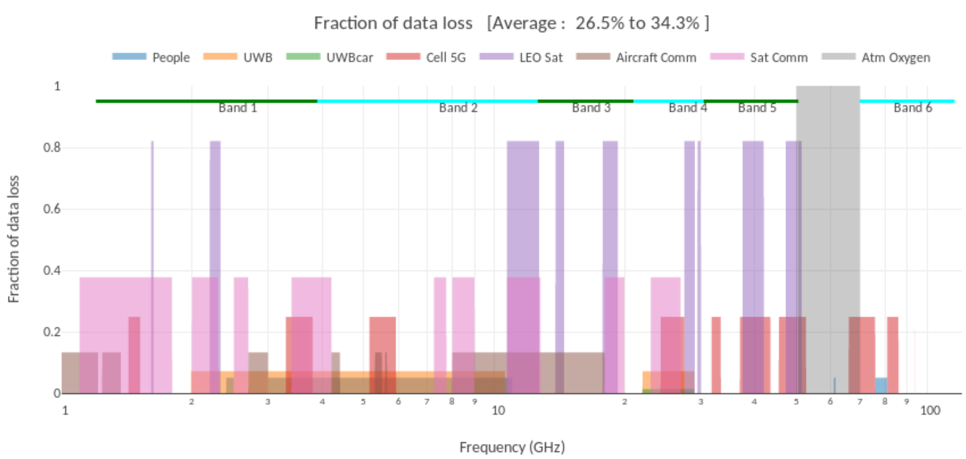
- [x] Post-Processing Flagging
- [ ] Antenna-based Real Time Flagging
- [ ] Baseline-based High time resolution Flagging (in-correlator)
- [ ] RFI modeling and subtraction at high time resolution

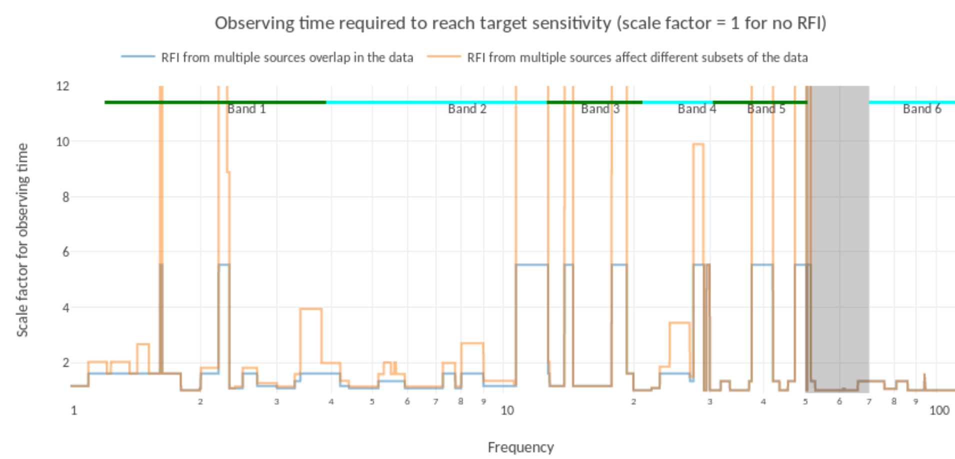
RFI Decorrelation

- ○ None (ignore from calculations)
- ◉ RFI at 20deg from phase-center (practical estimate)
- ○ RFI at 90deg from phase-center (maximal decorrelation)

Attenuation threshold

○ 20 dB ◉ 40 dB ○ 60 dB





https://gitlab.nrao.edu/rurvashi/ngvla-rfi-impact-simulator
https://github.com/urvashirau/ngVLA-RFI-impact-simulator

Conclusions

Status quo : ~ 30% data loss (on average). 100% in LEO-sat bands.
Observing time nearly doubles (from ideal) in order to compensate
No additional development/operating cost

High time-resolution outlier detection :

< 10% data loss in most bands (except cell 5G and LEO-sat)
No significant increase in required observing time (from ideal)
Additional development/operating cost

Regulation : Likely required for LEO-sat bands and 5G cell tower location where modeling and subtraction algorithms may work only for simple situations.

This entire analysis is based on many unverified assumptions

- Does the RFI have gaps at high time-resolution ?
- How many interferers can be modeled and subtracted together ?
- How much does decorrelation help/hurt ?
- How viable is the idea of an integrated smart database ?

=> Construct prototypes + borrow from experiments at other observatories.