

A cautionary tale about pipelines

It's time to raise the bar

Frits Sweijen

LOFAR Family Meeting, 04/06/2024
Surveys & Methods

Table of contents

Why this talk

Data reduction: LINC Target HBA DI selfcal

“De rode draad”: we need better quality control in pipelines going forward

Examples that should have failed QC

Why this talk?

Updates to LINC HBA

DI selfcal for difficult fields like galactic plane, bright sources or complex diffuse emission

...but these serve a greater narrative

Lots of knowledge gathered over the past decade or so

LoLSS, LoTSS, full ILT: this should flow back into the (base) pipelines

Workflows have gotten more complicated, goals more ambitious

Direction-dependent calibration, polarisation, long baselines: strict quality control will make life easier

LOFAR2.0 programs will only make this worse

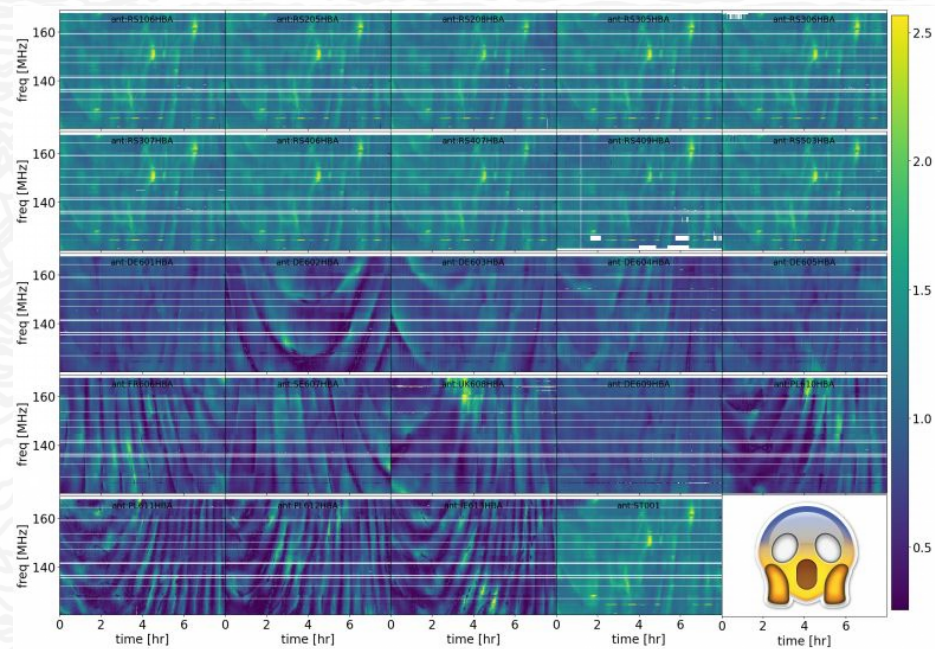
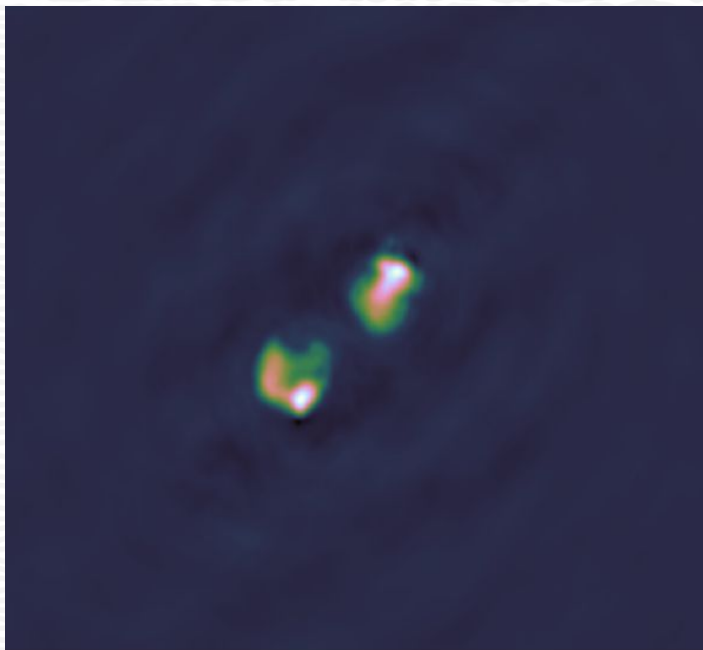
e.g. iLoTSS, LLoCuSS, LUDO

Data deletion protocols demand “sustainable calibration”

No chance to go back and improve later, so we can't compromise early on

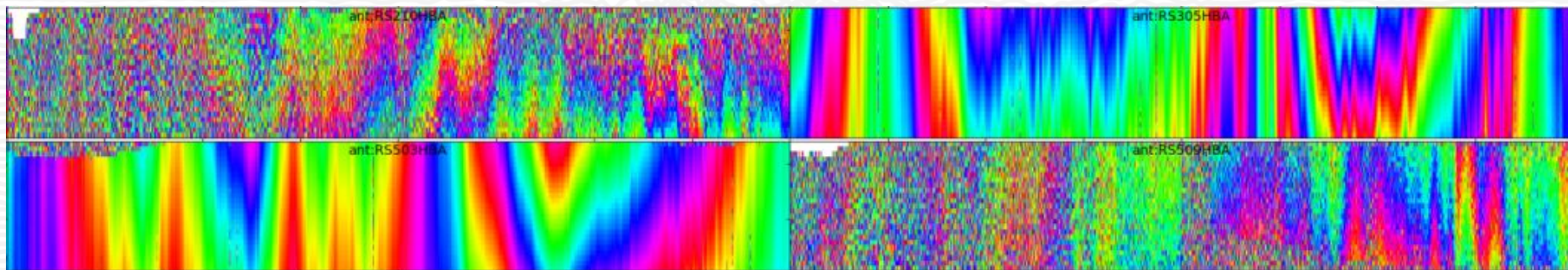
How it began

ILT calibration of 3C 295 not quite working



ILT calibration of 3C 295 not quite working

Root cause: unacceptable solutions on Dutch stations!



Re-evaluating the LINC HBA target strategy

Current strategy ignores known limitations

We know the TGSS model is incomplete for LOFAR purposes

Limitations of TGSS

1. Lower resolution: 25" vs 6"
2. Lower sensitivity: 3.5 mJy/b vs 0.095 mJy/b median rms noise
3. TGSS ADR1 has 200λ inner uv-cut
LOFAR goes down to $\sim 34\lambda$

Limitations of strategy

1. High time resolution, high frequency resolution, XX and YY separately
2. Discrete solutions between frequency chunks

Many degrees of freedom in an under-constrained problem!

*Mistakes made in this way
can not be corrected later*

Updates to LINC HBA

OLD STRATEGY

1. Split 48 MHz bandwidth in 24 independent 2 MHz chunks
2. Only one calibration against TGSS model
3. Diagonal: XX and YY phase solutions

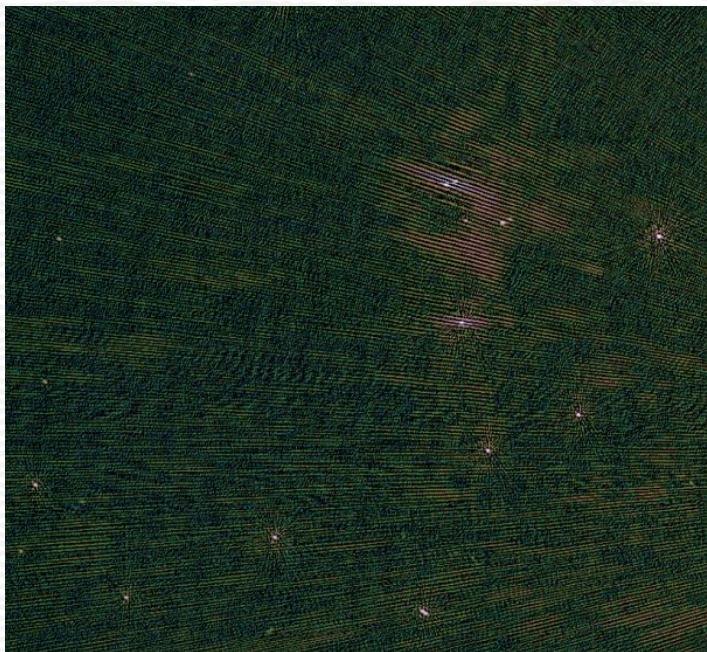
UPDATED STRATEGY

1. Treat full bandwidth as 1x 48 MHz chunk with physically motivated constraints
2. First calibration using TGSS model
3. Image full FoV at 6" resolution
4. Second calibration against LOFAR model
5. Scalar: polarisation independent phase solutions

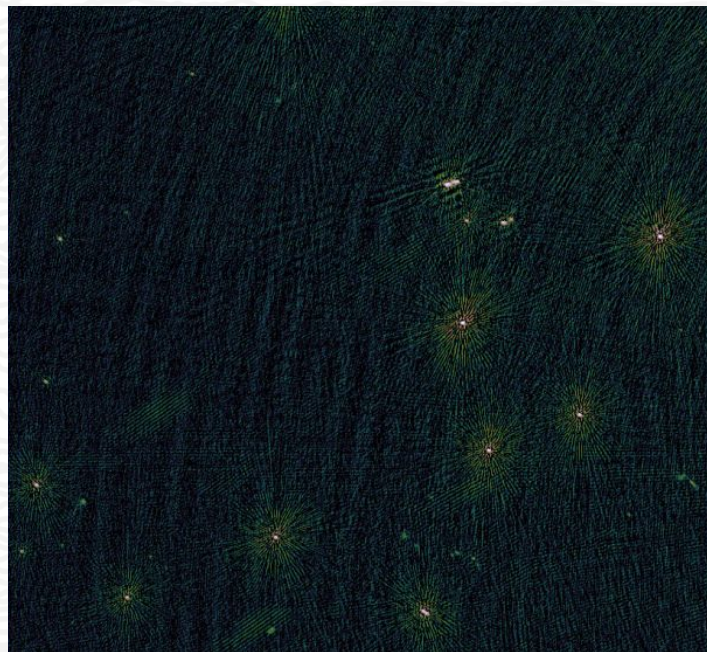
Better future improvement: start from LoTSS?

Image plane consequences

6" DI image - old strategy



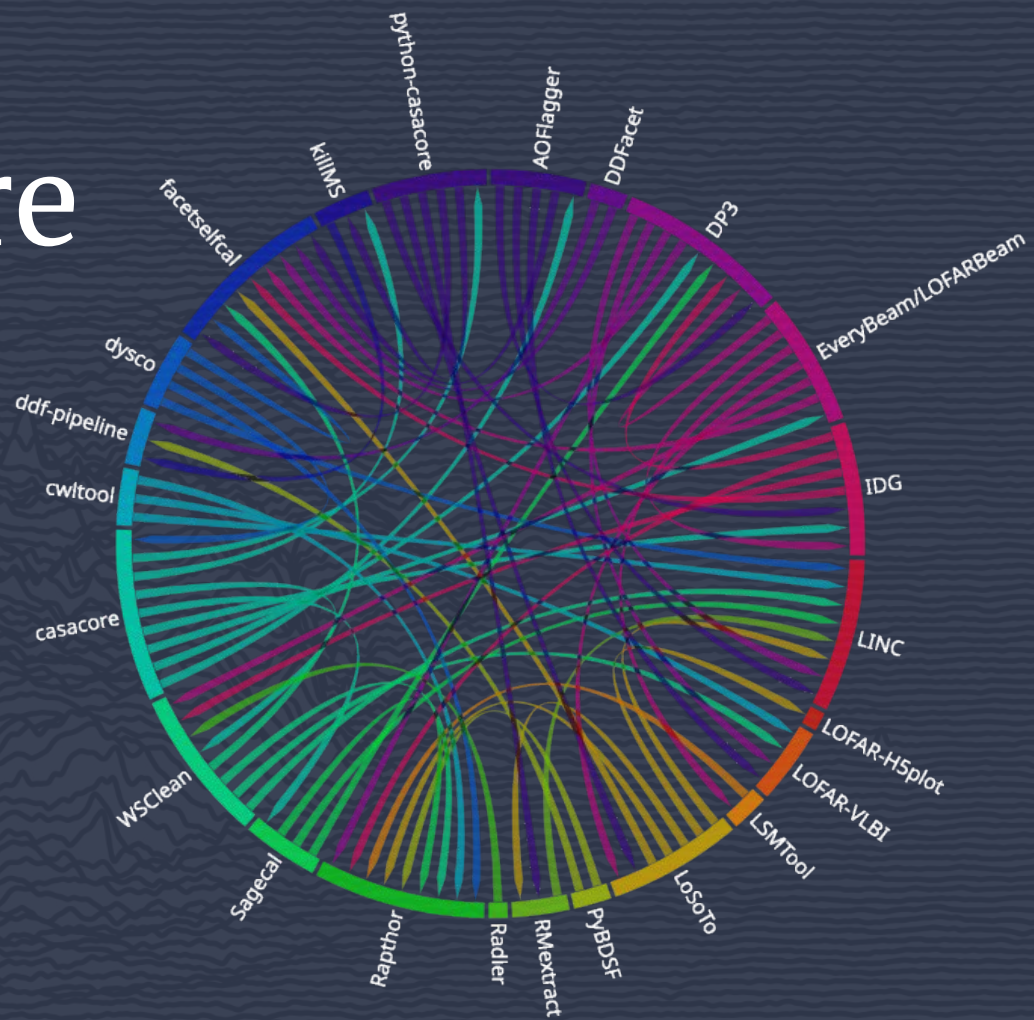
6" DI image - new strategy



Now for the greater
narrative

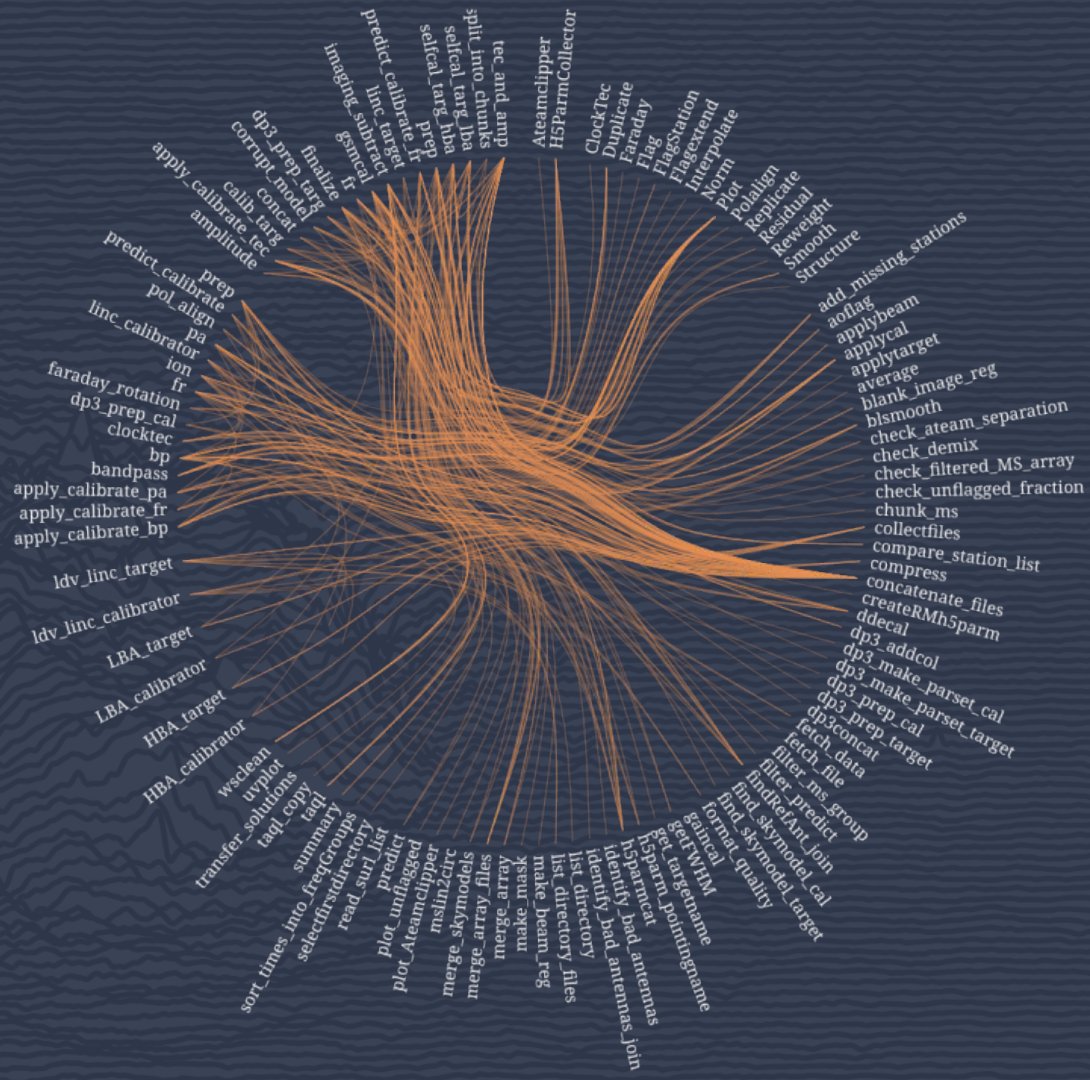
Many software packages

Lots of dependencies



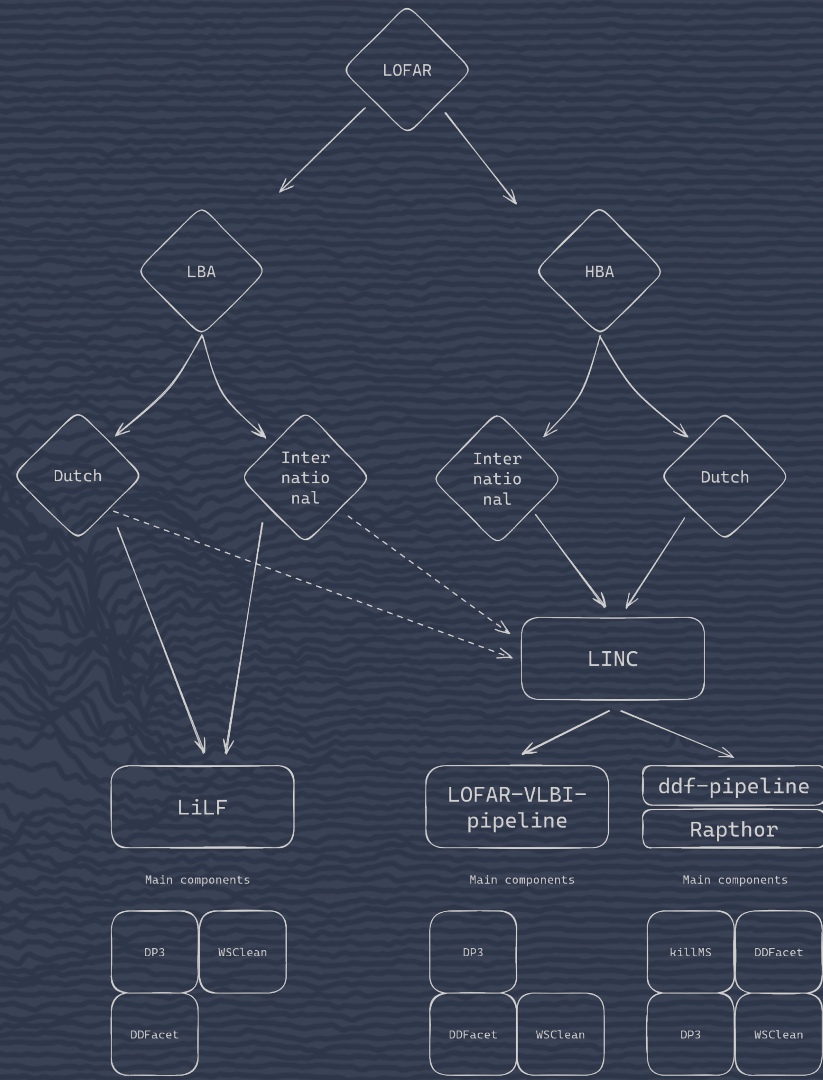
Complicated pipelines

Hundreds of inputs, dozens of steps



Moving beyond “the pipeline”

It is time to get picky about our data and reflect on strategies



Moving beyond “the pipeline”

The pipeline itself is no longer the point or goal

Reflect on possible updates to trickle down

Strategies “out of date”? VLBI-suitable pre-processing in LINC? etc.

User friendliness

Limit the number of required inputs

Pipelines are complicated, running them shouldn't be

Be strict about quality control e.g. don't want to wonder/worry about LINC when doing VLBI

Indicate (potential) problems to the end user

When to abandon the data?

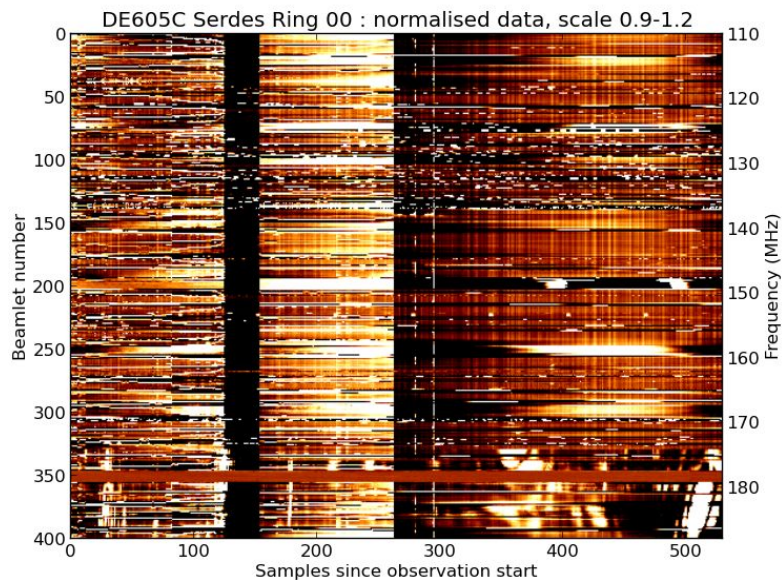
Fail (gracefully) on data or calibration not meeting criteria

Pipelines create a sense of trust, irrespective of how well they work

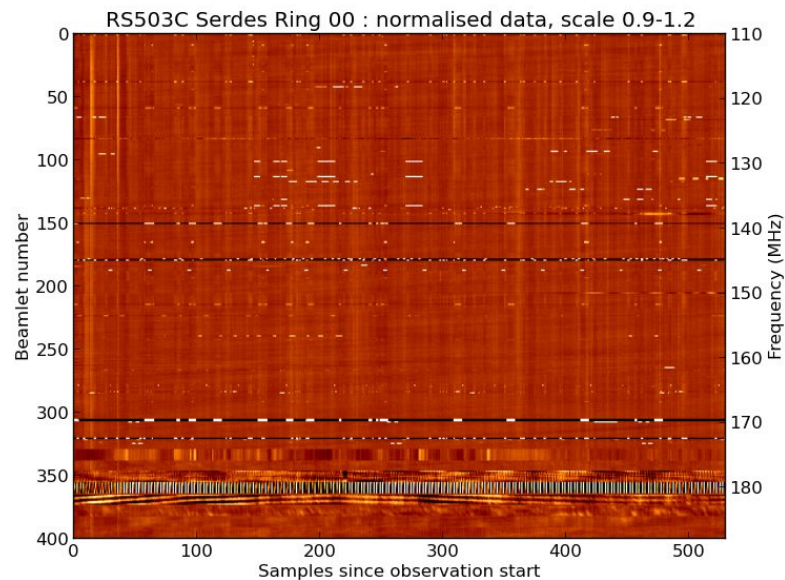
Visual quality assessment
is subjective and hard!

Inspection plots after observing

Obviously bad

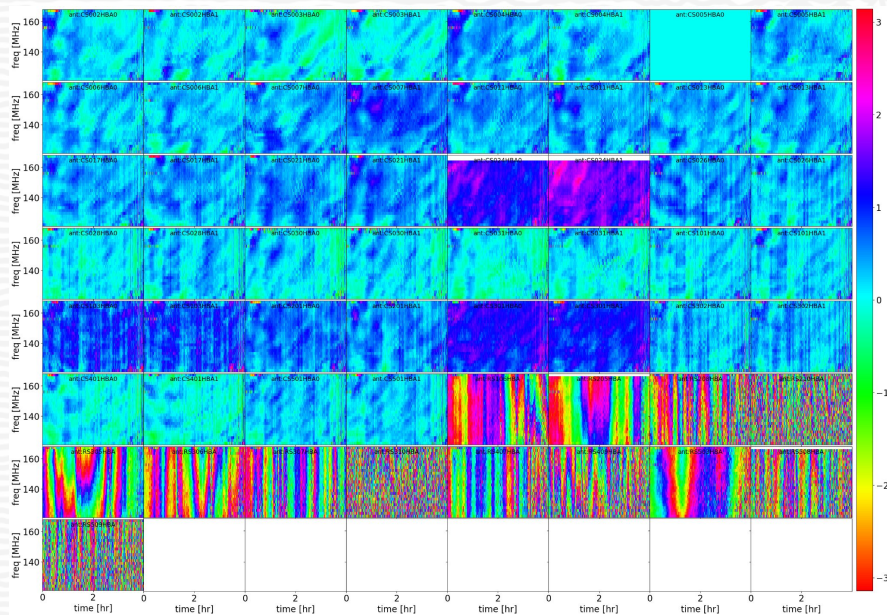


Looks "fine", but did not survive LINC!



Calibration solutions from LINC

We currently accept/ignore this



While we could have this

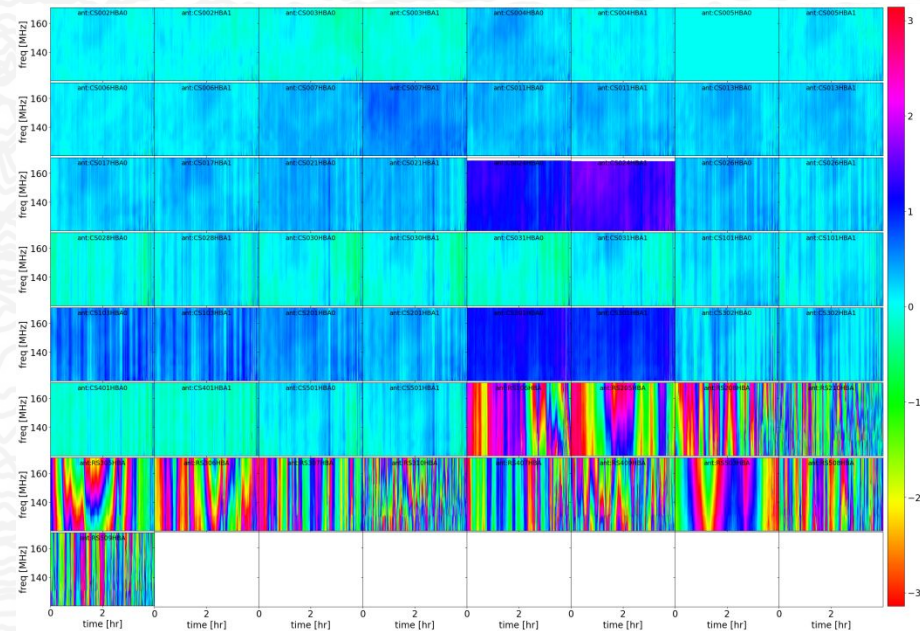
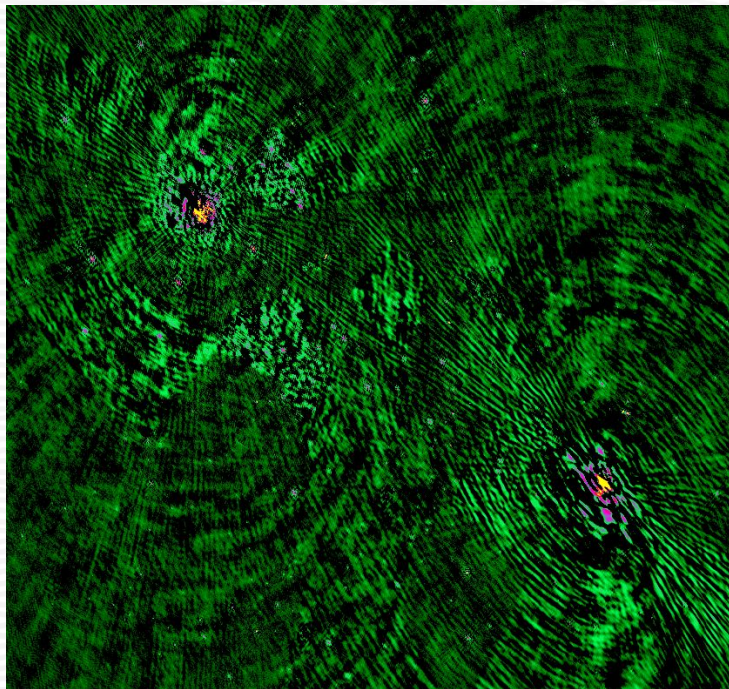
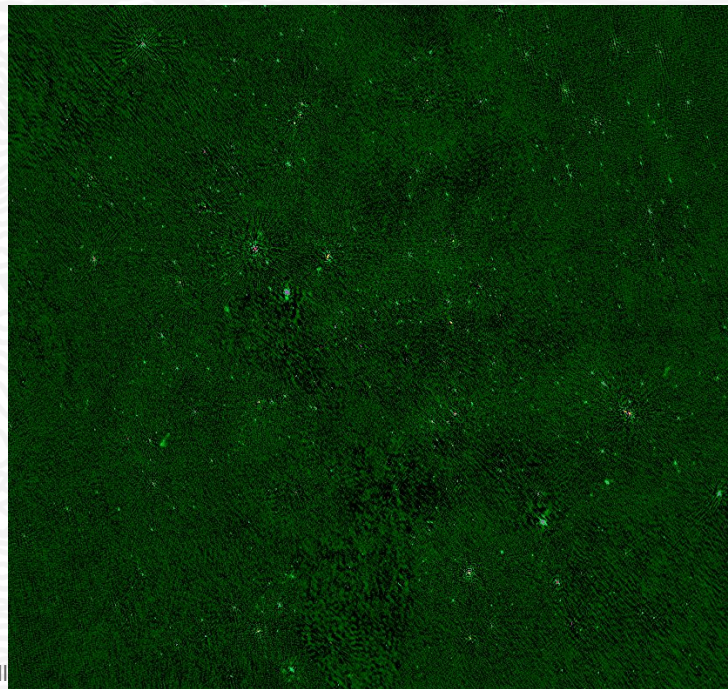


Image plane consequences

Old strategy



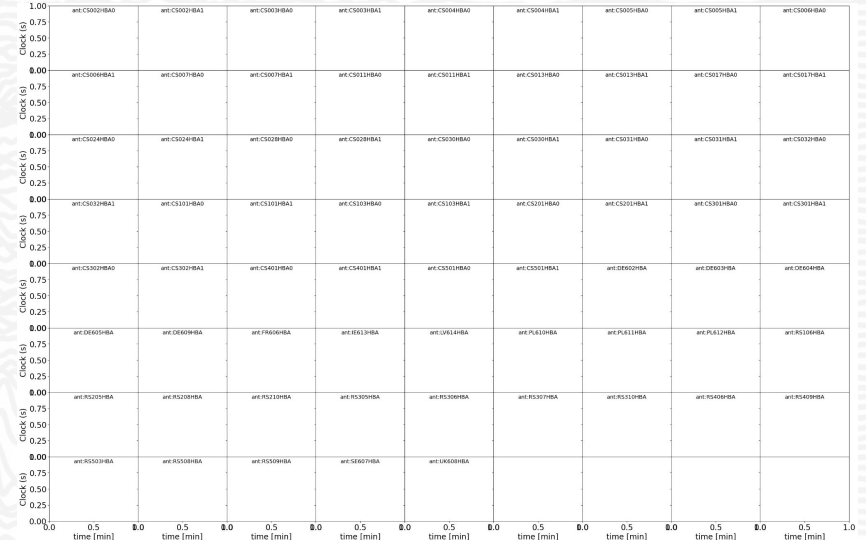
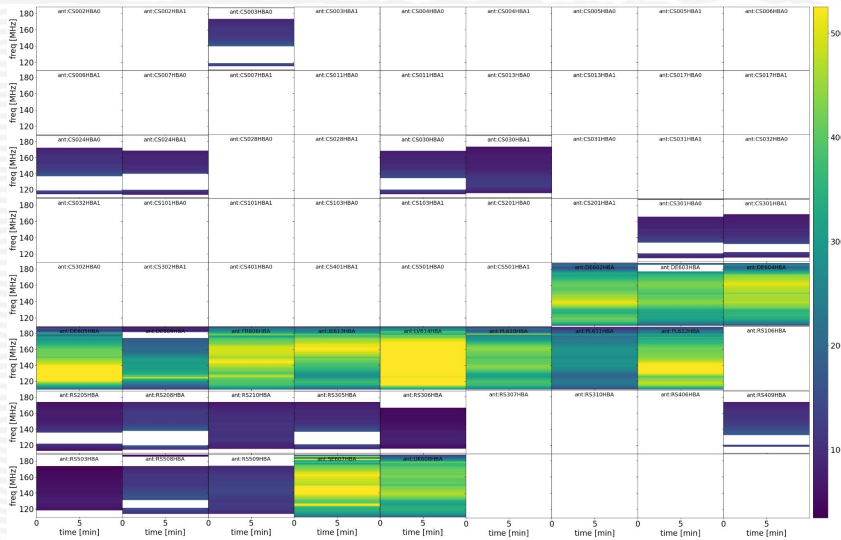
New strategy



Courtesy of T. Shimwell

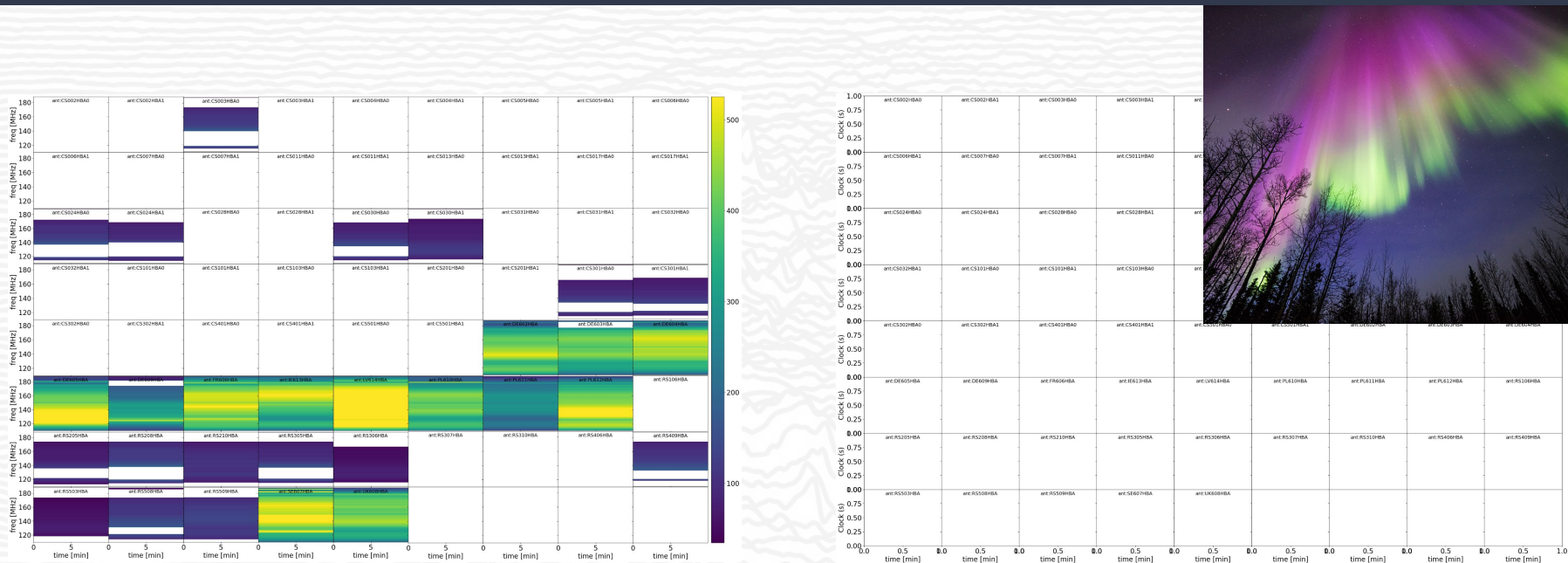
But not always

“INFO: Final process status is success” .. or is it



Nothing crashed, so it ran “successfully”, but nothing is usable either, i.e. it should have failed.

“INFO: Final process status is success” .. or is it



Nothing crashed, so it ran “successfully”, but nothing is usable either, i.e. it should have failed.

Closing off

Thank you developers and maintainers for making complaints possible

Software is hard; complaints are a sign that things are working and people are interested

Say no to status quo

- TGSS is not sufficient, DI selfcal in LINC should be the new default for HBA (as already done for a long with with e.g. ddf-pipeline)

Too much problematic data is allowed through with bandaids in later pipelines

- “No data is better than bad data”, cut your losses and reobserve

Pipelines need to get more robust and automated quality control, especially the early ones like LINC

- Inspecting data and calibration is time consuming and tricky for both experts and non-experts

Lots of experts out in the field

- Community work are important signposts where to spend effort or what the pitfalls are
- Regular barometer needed to keep efforts and desires aligned

Thank you!