LOFAR2.0 processing pipelines and associated commissioning activities

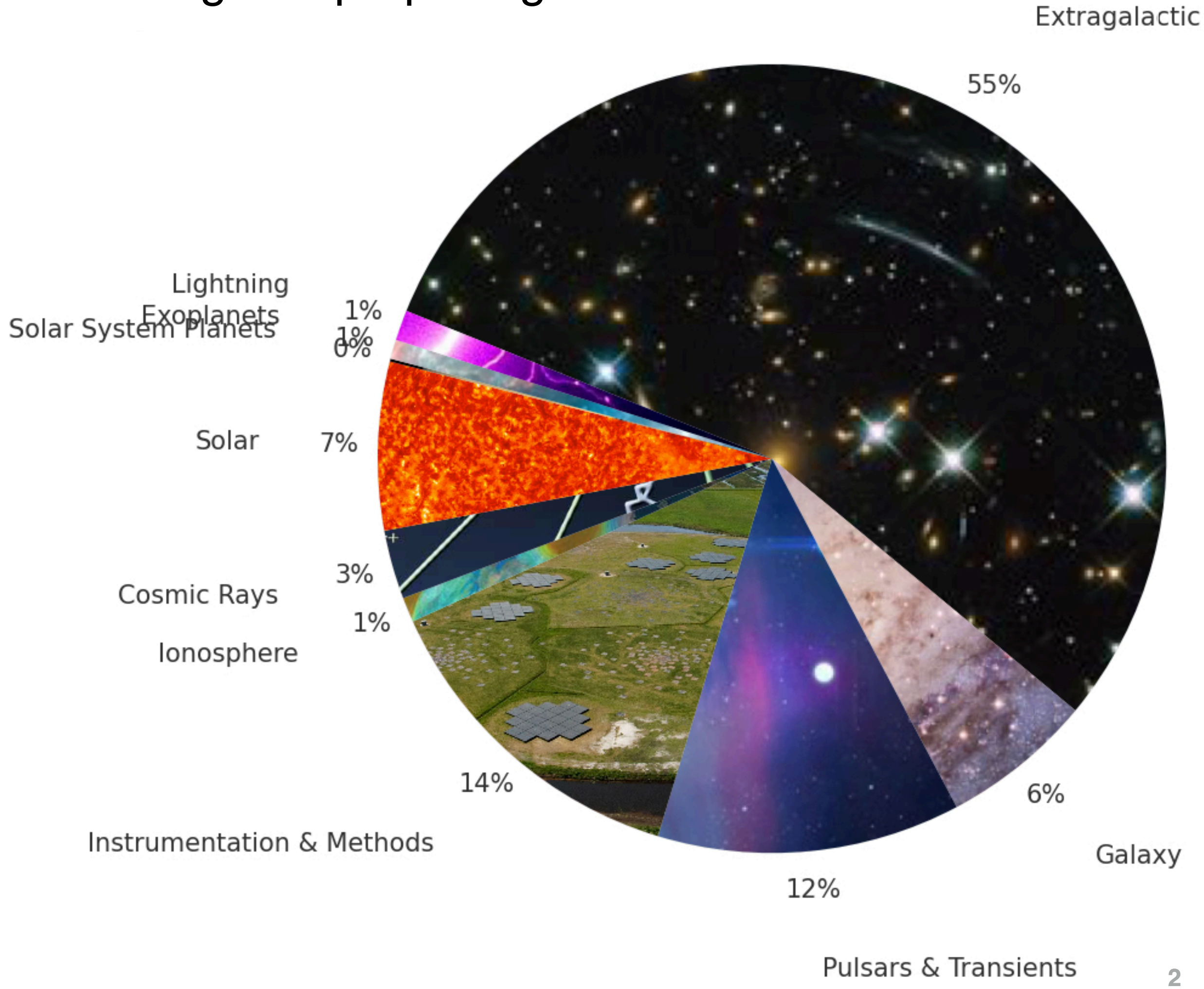
Towards LOFAR2.0 science

Almost 1,000 publications from LOFAR over a wide range of science areas — one of the best value astronomy facilities and amongst the most impactful at radio wavelengths.

LOFAR2.0 increases the scientific potential with enhanced telescope performance and more streamlined processing, observing and proposing.

Community is scientifically ready for LOFAR2.0 with 15 large proposals (over 17,000hrs) involving 500+ researchers.

New users may be attracted if science-ready products can be provided in various areas.



Towards LOFAR2.0 science

Existing pipelines allowed science in all areas covered by the LOFAR2.0 large programs.

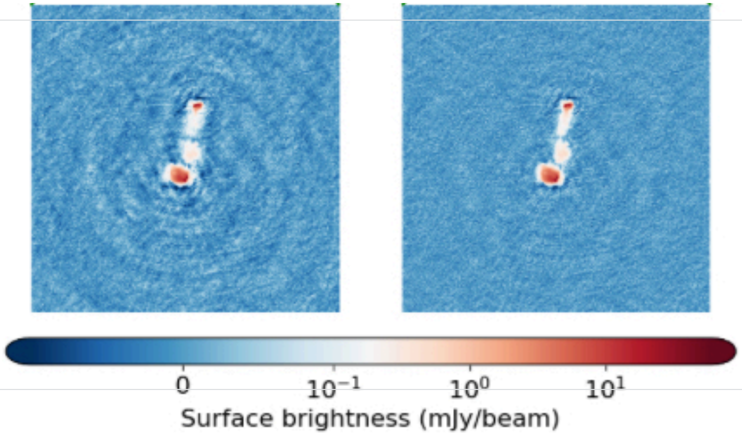
However, in many cases pipelines were not ready when LOFAR started and there was a delay in scientific output. A very different position for LOFAR2.0 thanks to the development over the last 15 years and the extensive efforts presently underway.

Of course still some cause for concern:

- LOFAR2.0 can observe pretty quickly
- Storing large datasets is expensive.
- The LOFAR2.0 programs are quite ambitious.
- The computational expense of some of the programs may remain very high.
- Continued research and development required to optimise efficiency and automation of pipelines.
- In some cases the quality may remain suboptimal

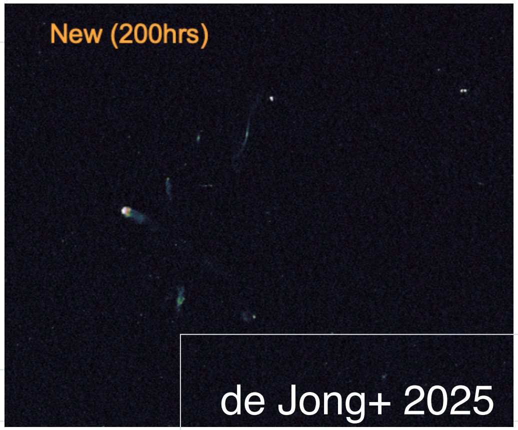
Towards LOFAR2.0 science

Progress has been excellent in many areas. For example, in the HBA VLBI case (the most computationally expensive projects) a very active community has dramatically enhanced our capacity to produce images at 0.3” resolution.



Despite reduced CPU cost (~15x decrease in 4 years) quality has also improved

Depth now 1/3 of the deepest LOFAR2.0 ambitions



de Jong+ 2025

Some VLBI busy weeks.

but computational cost of LOFAR2.0 programs remains enormous. The largest LOFAR2.0 projects (~1/4 billion core hours) will be challenging to fully realise.

Towards LOFAR2.0 science

Lots of pipelines/tools for different science aims. The list that may be needed for LOFAR2.0 large programs includes:

Pre-processing	**LINC**	**DDF-pipeline**	**RAPTHOR**
LOFAR VLBI pipeline	**LiLF**	**Facetselfcal**	**TrAP**
Extraction	**DynSpecMS**	**PulP**	**AARTFAAC EoR**
Lightning Impulsive imaging	**Lightning Beamformed imaging**	**Cosmic Ray**	**LORDS**
RMsynthesis	**PyBDSF**	**Solar**	

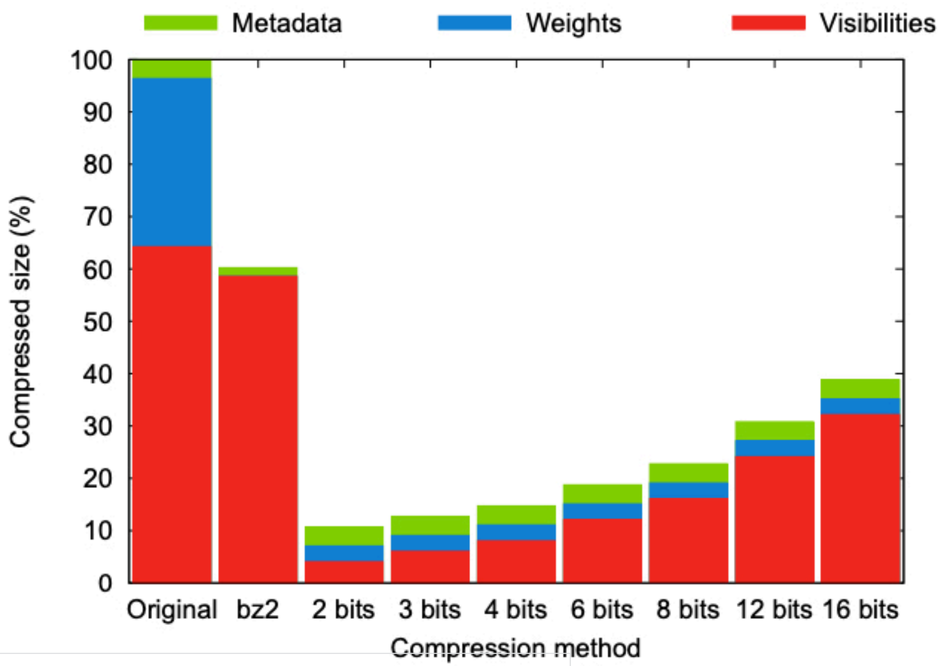
Pre-processing pipeline

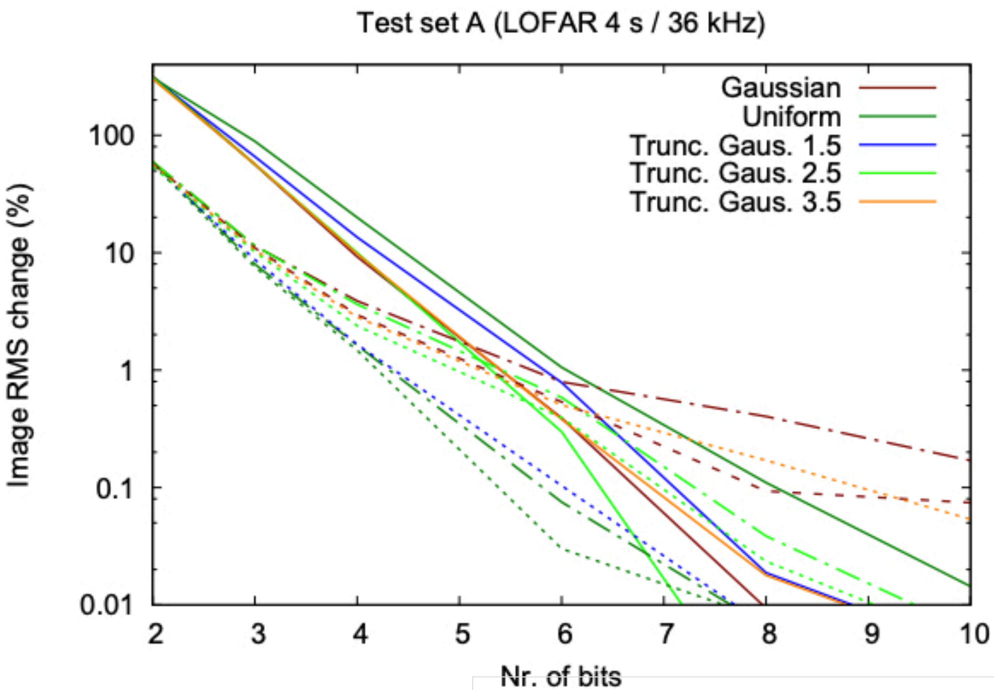
https://lofar-pre-processing-pipeline.readthedocs.io/en/latest/

The same intentions as LOFAR1.0 (i.e. optionally flag, demix, average, compress) for HBA and LBA imaging data on CEP6.

Changes include:

- Now in CWL and more closely coupled with the initial LINC stages.
- Demix strategy revised (also for LINC).
- Ongoing investigation into optimal dysco compression settings (maybe ~30%) and loss-less metadata compression.
- Production of quality assessment products and automatic identification of failed observations





- Slide adapted from M. Iacobelli.

- Plots from Offringa 2016

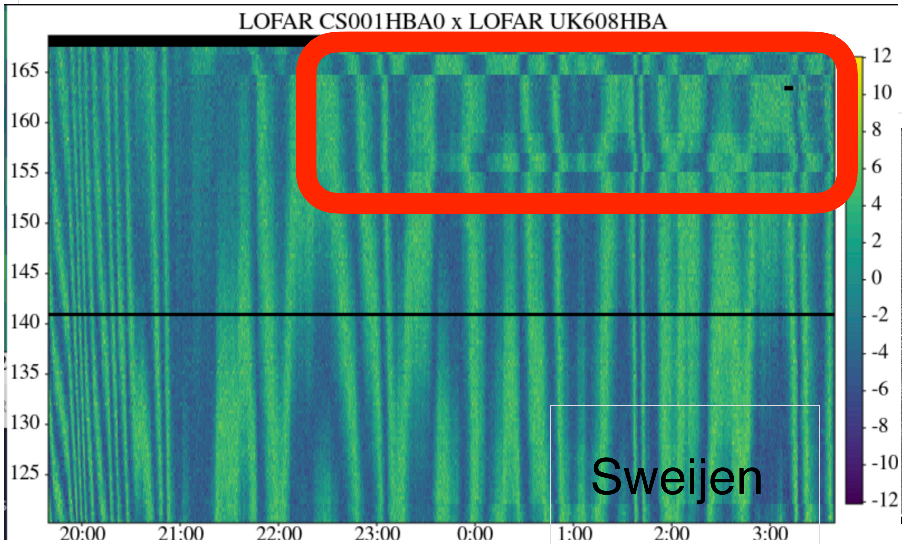
LINC

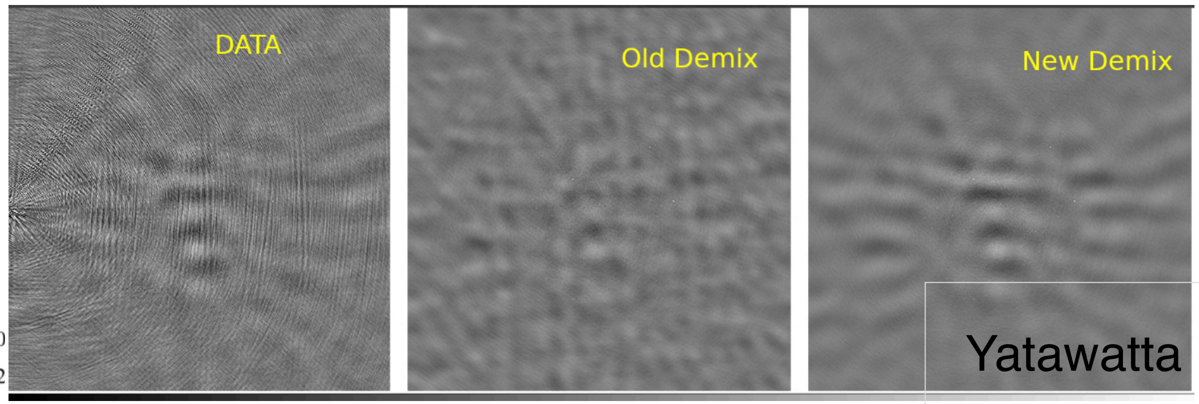
https://linc.readthedocs.io/en/latest/

Initial direction independent calibration of the target field data (mainly HBA but LBA functionality also exists).

Main changes in the past months include:

- Now also optionally outputs data ready for the VLBI pipeline
- Smooth full bandwidth solve using improved sky models.
- Faster A-team clipping.
- Demix strategy revised.
- Move from RMextract (no longer maintained) to Spinifex to rotation measure corrections.





- Slide adapted from team RAPTHOR.

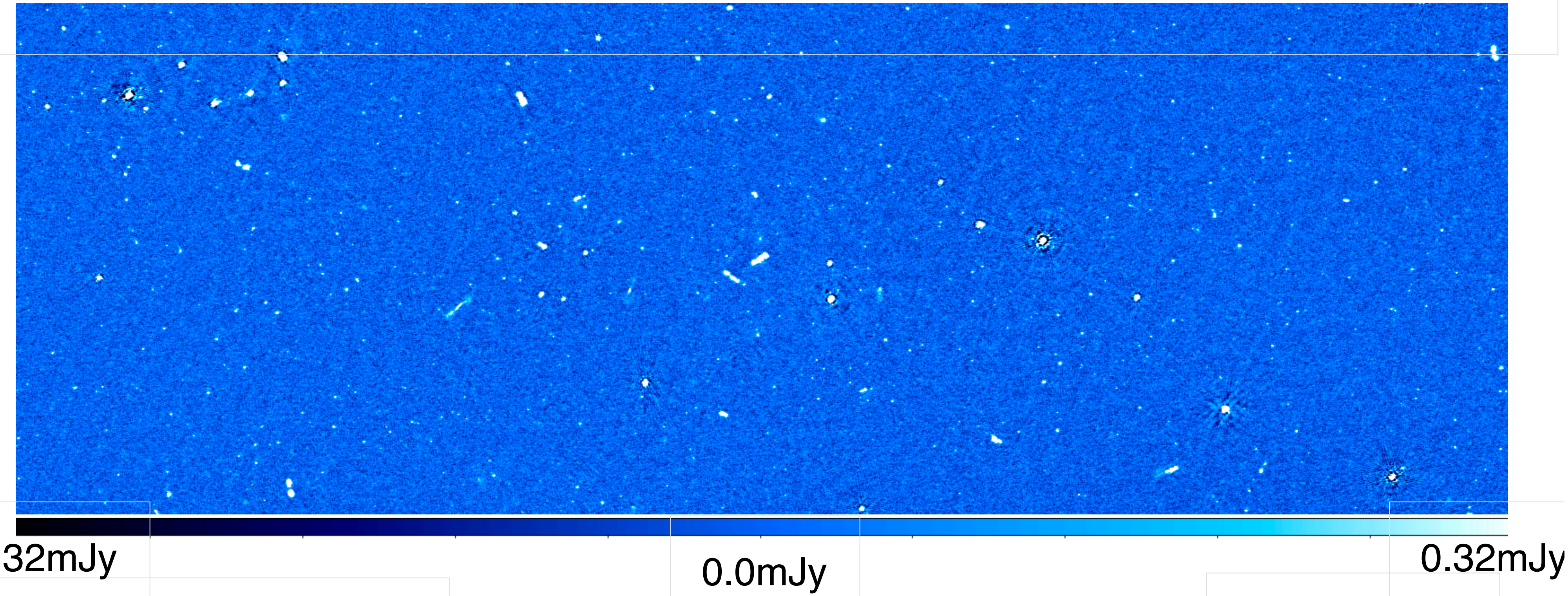
RAPTHOR

https://rapthor.readthedocs.io/en/latest/

Direction dependent calibration and imaging of the target field data (HBA) produces Stokes I, Q, U, V image products and calibrated data.

Recent improvements include:

- Image quality (refined strategy and component improvements)
- Speed (due to component improvements)
- Stability (self calibration convergence checks)
- Refinement of flux density scale and astrometry



- Slide adapted from team RAPTHOR.

- D. Rafferty

DDF-pipeline

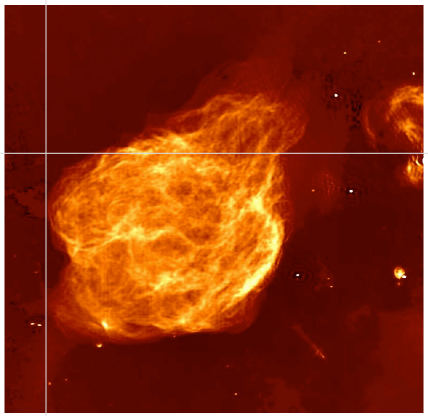
https://github.com/mhardcastle/ddf-pipeline

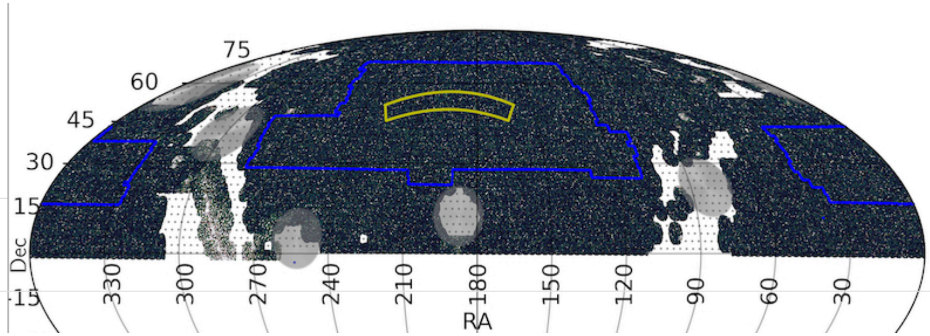
Direction dependent calibration of the HBA produces Stokes IQUV image products, dynamic spectra and calibrated data. Characterised product quality and performance (75hrs per 8hrs data) using 15,000hrs of data covering 90% of sky through ~400 publications.

Ongoing development on:

- More constraining options and faster solving
- Speed improvements (2x faster calibration and GPU gridding/degridding)
- More speed improvements (initial tests with an MPI version indicate an ~24 times speed up when using 24 cores)
- Improved deconvolution (mosaicing, RM deconvolution, better deconvolution of large scale structures)
- Multi-field option to calibration and imaging just postage stamp regions (VLBI applications)
- Refined post processing pipelines (transients, polarisation, dynspecms and extract/selfcal).

Example model image of galactic structure. C. Tasse





The 89% of the sky imaged with DDF-pipeline

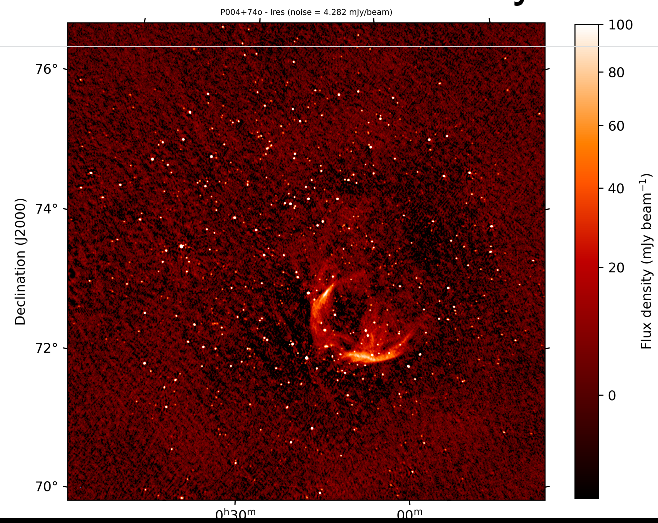
Slide adapted from Tasse, Hardcastle, Shimwell

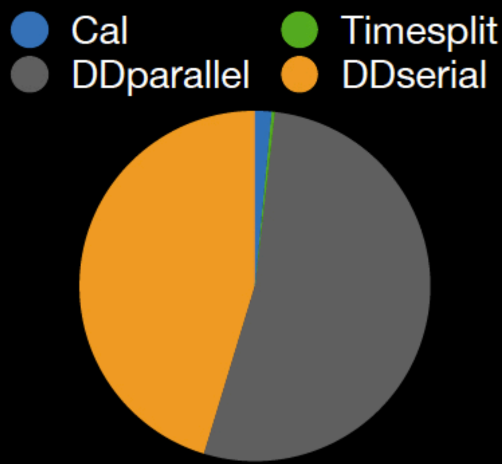
LiLF

Performs direction dependent calibration of LBA imaging data and produces Stokes IQUV image products as well as calibrated data. LiLF was developed for the LoLSS survey.

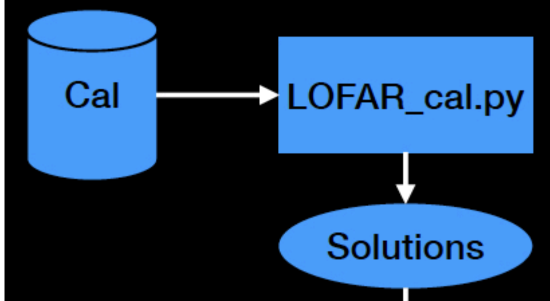
Recent changes include:

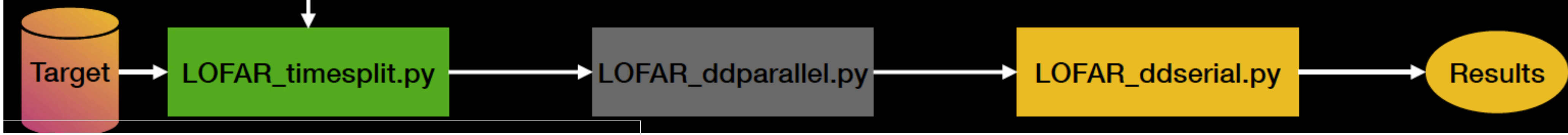
- Rewrite of the pipeline with revised calibration strategy.
- Significant improvement in quality since LoLSS-DR1
- Extensive characterisation (so far 40+% of LoLSS)
- Ongoing efforts to provide detailed documentation and prepare for LOFAR2.0 operations





Pipeline	CPU hours
Cal	100 /2
Timesplit	20 /2
DDparallel	3500 /2
DDserial	3000 /2





- Slide adapted from F. De Gasperin & H. Edler

VLBI

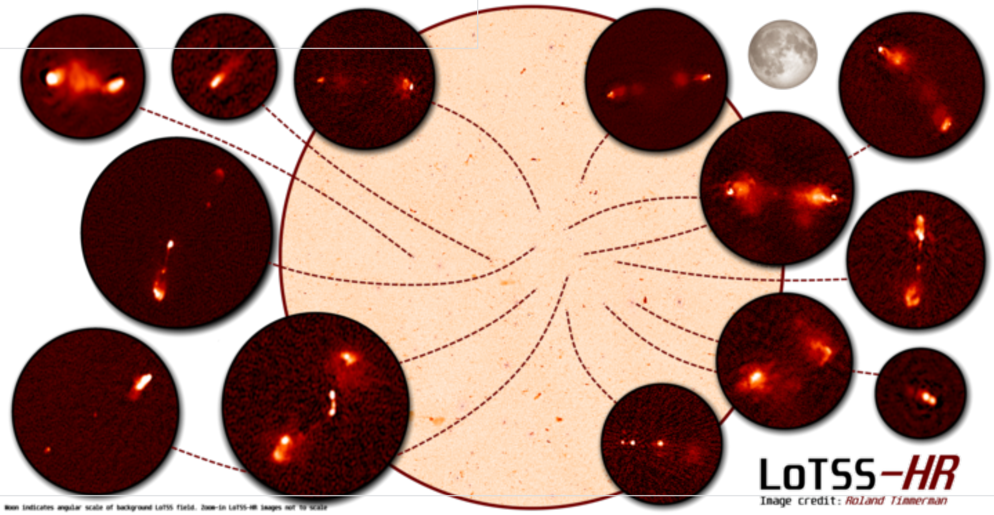
https://github.com/LOFAR-VLBI/lofar-vlbi-pipeline/wiki

High-resolution (up to 0.3”) imaging using the full ILT. In the final stages of completing workflow and aim to have end-to-end non-interactive run by end of September. Will produce postage stamp and/or wide field images.

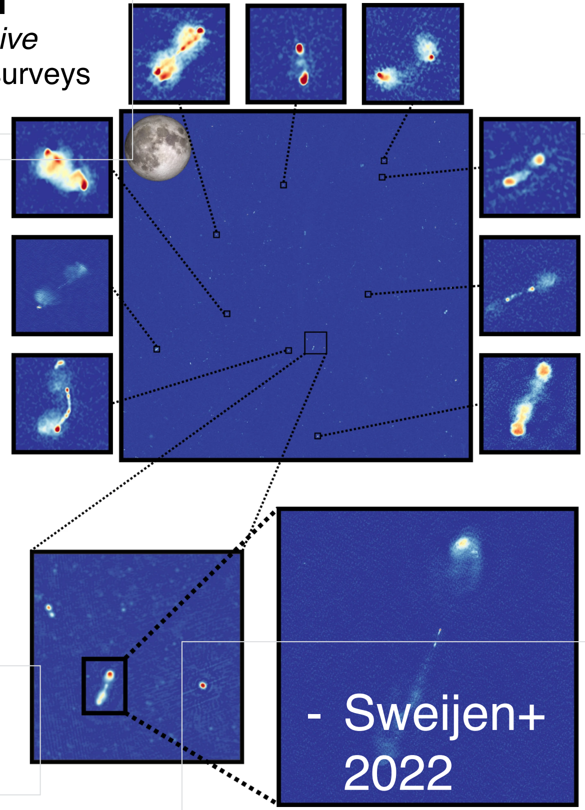
Ongoing activities include:

- Detailed profiling of the entire process (presently ~50k per 8hr field)
- Incorporating polarisation
- Extending to LBA

LOFAR-VLBI pipeline
(relatively) computationally cheap
Application: wide area surveys



Widefield LOFAR-VLBI
computationally expensive
Application: deep field surveys



- Sweijen+ 2022

Join the VLBI working group mailing list - email leah.k.morabito@durham.ac.uk

Slide adapted from L. Morabito

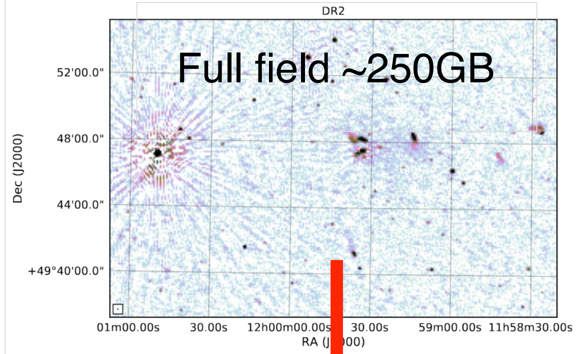
Facetselfcal

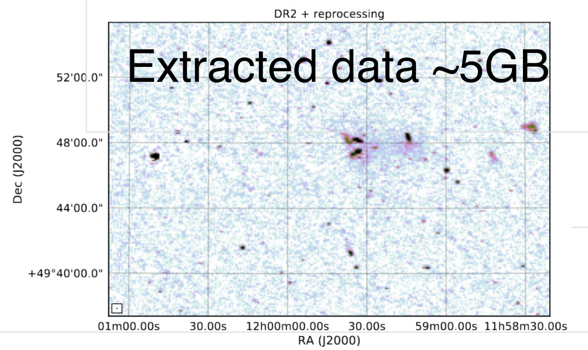
https://github.com/rvweeren/lofar_facet_selfcal

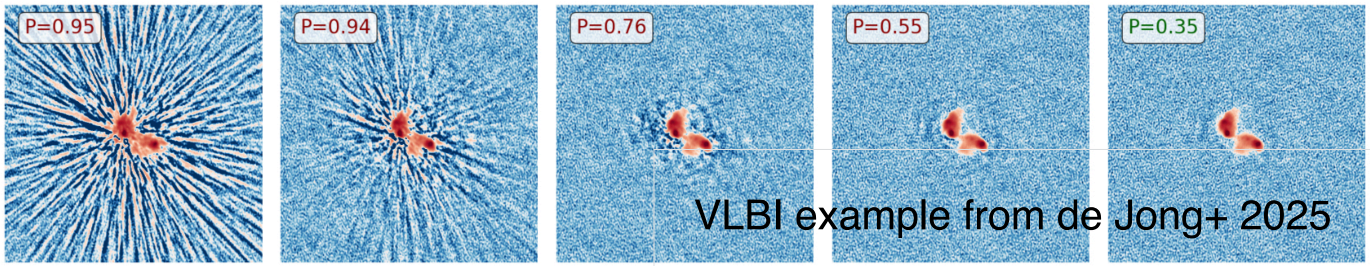
Performs direction dependent or independent calibration of HBA or LBA imaging data. Routinely used for extract targets of interest and tune calibration in NL-only imaging. Also to perform the self calibration in the VLBI pipeline.

Recent improvements include:

- Improved speed (Stokes I data only)
- Improve stability, quality and automation (tunes calibration parameters and identifies when self calibration is complete)
- Improved functionality (direction dependent options to study larger objects)





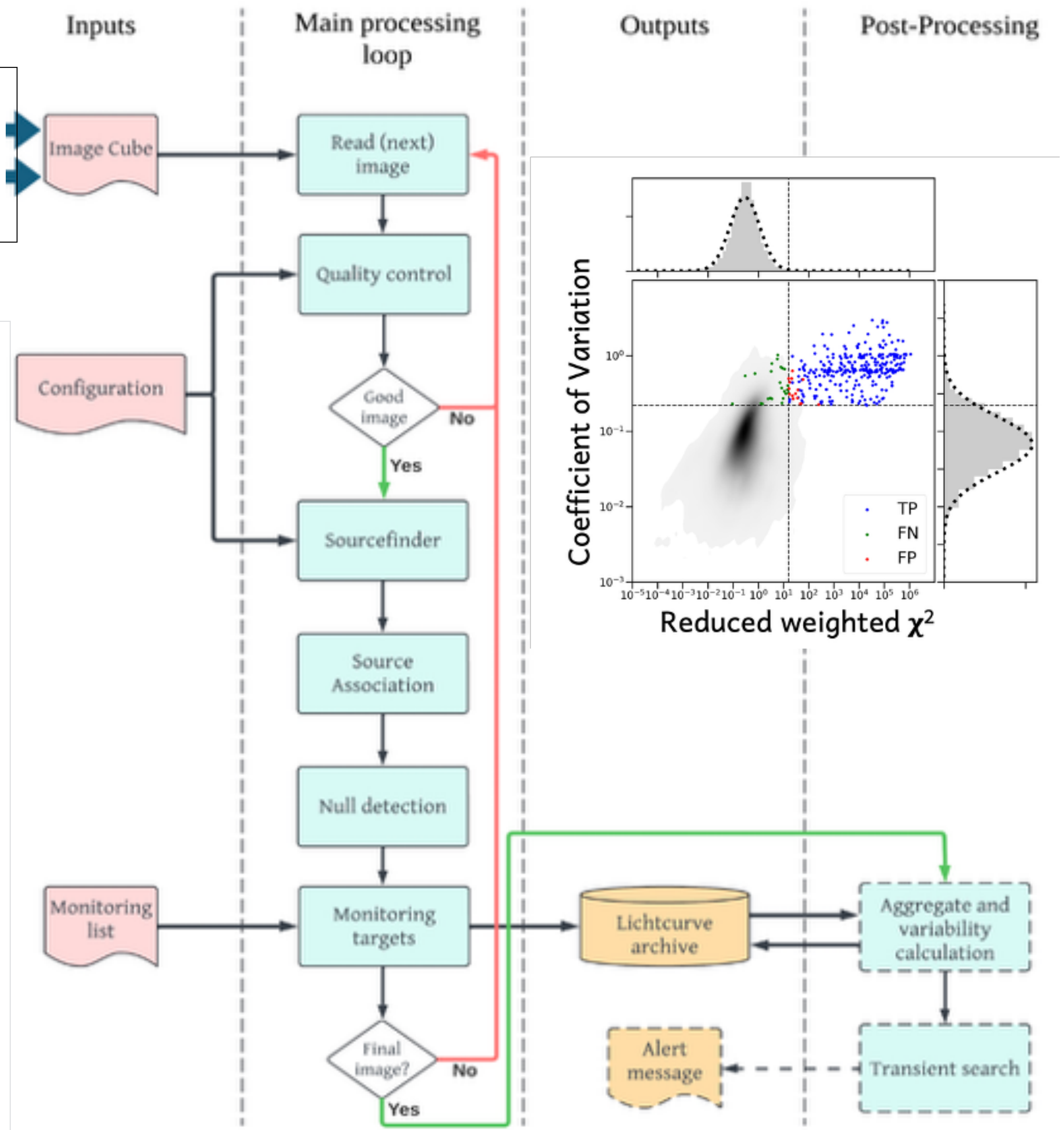


VLBI example from de Jong+ 2025

Facetselfcal: van Weeren+ 2021

Trap2.0

LINC full bandwidth and snapshot images



The flowchart of the LOFAR Transients Pipeline 2.0 (TraP2.0). An SDC-ready automated pipeline based on TraP (Swinbank et al. 2015).

Automated pipeline for radio transient searching based on TraP (Swinbank+ 2015). Intends to search for transients using LINC products (deep images and source-subtracted snapshot images).

Python Software rewritten to improve efficiency (remove calculations and now multithreaded) and runtime now ~20 times faster.

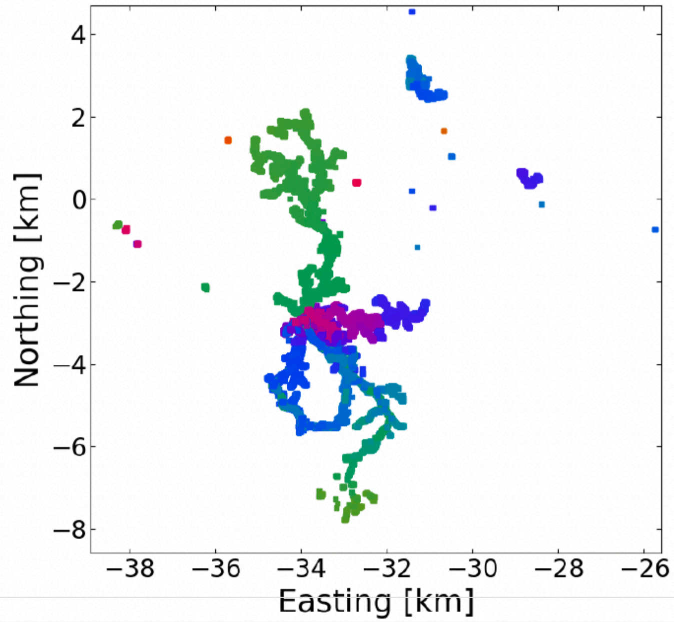
Still to work on:

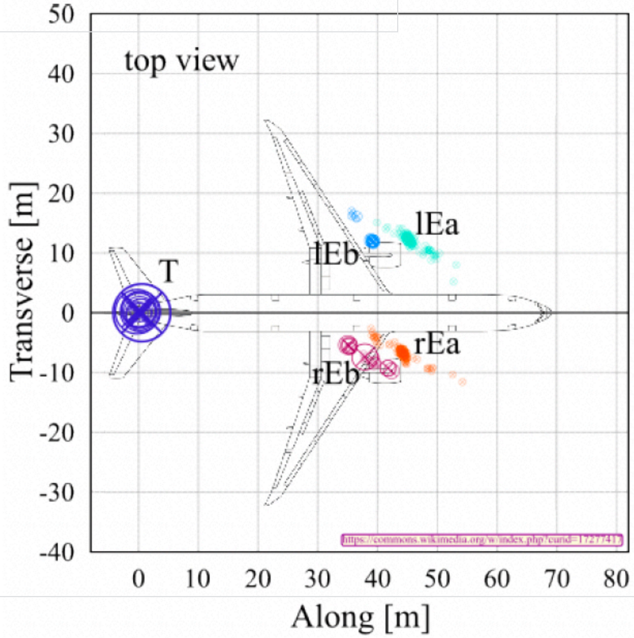
- missing multi-frequency processing
- Not yet in CWL or alternative.

Slide adapted from A. Rowlinson and team RAPTHOR

Lightning

Two imaging types which produce world-leading lightning images





Impulsive imager (time of arrival based) images the entire flash

Nearfield beam forming imager is higher resolution but can only image small area (~100m boxes).

Imagers used for LOFAR1.0 will be reused for LOFAR2.0:

- The imagers use transient buffer data that has been cleaned of RFI and undergone quality control. Calibration is implied by the single LOFAR2.0 clock.
- The impulsive imager will be automatically run to image the entire flash.
- The near field imager will be used to obtain more detailed images where required.
- Work needed to ensure compatibility with LOFAR2.0 data and on automated framework for the impulsive imager.

Slide adapted from B. Hare

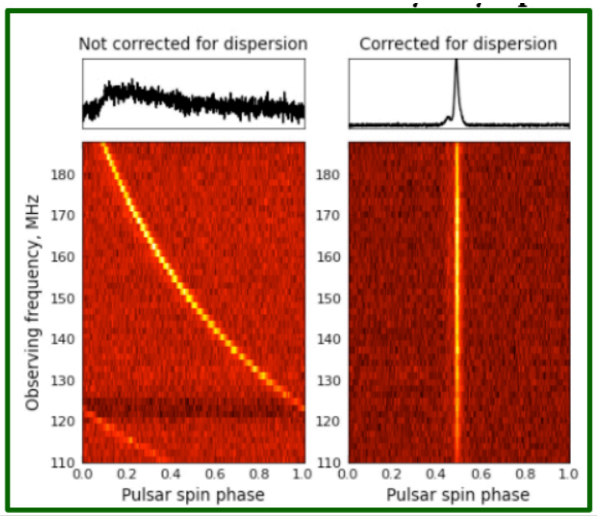
Pulsar Folding Pipeline (PulP2)

https://lofar2-beamformed-pulsar-folding-pipeline.readthedocs.io/en/latest/

New pulsar folding pipeline (PulP2) is the CWL implementation of the original PulP (Kondratiev+ 2016) pipeline. The aim is to get the average profile of pulsars and provide data cubes for further analysis (it is not a search pipeline).

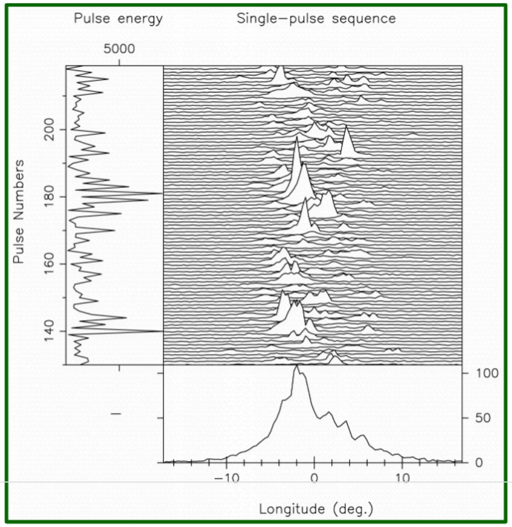
PuLP2:

- Read in raw 8-bit or 32-bit HDF5 complex voltage beamformed (XXYY) or Stokes I data and input parameters
- Using DSPSR (van Straten & Bailes 2011) the data are coherently dedispersed and folded according to the supplied pulsar timing model
- Optionally removes RFI and optimise the pulsar period and dispersion measure.
- Outputs diagnostics, logs, summary files as well as folded data (freq, time, pulse phase, polarisation cubes) in a standard format.



Credit: Anya Bilous

Dedispersion (correcting ISM delay)



Folding to align the pulsar rotations

Slide adapted from Vlad Kondratiev

Solar imaging

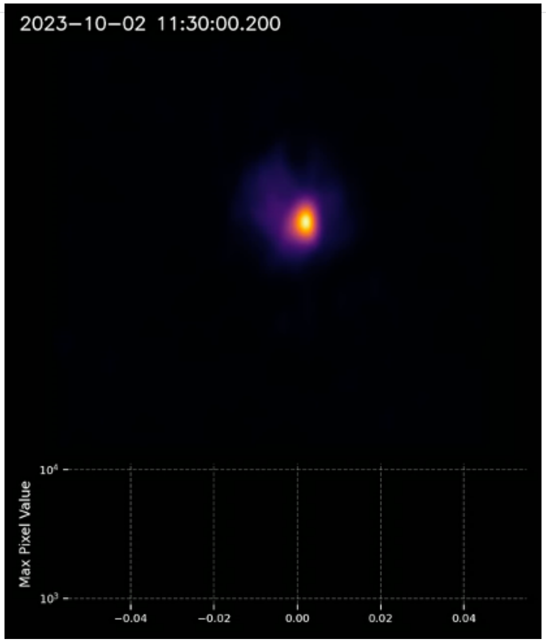
New solar imaging pipeline (SIMPL) described in Dey+ 2025. Tackles challenges due to rapid, several order of magnitude, brightness variations as well as ionospheric distortions.

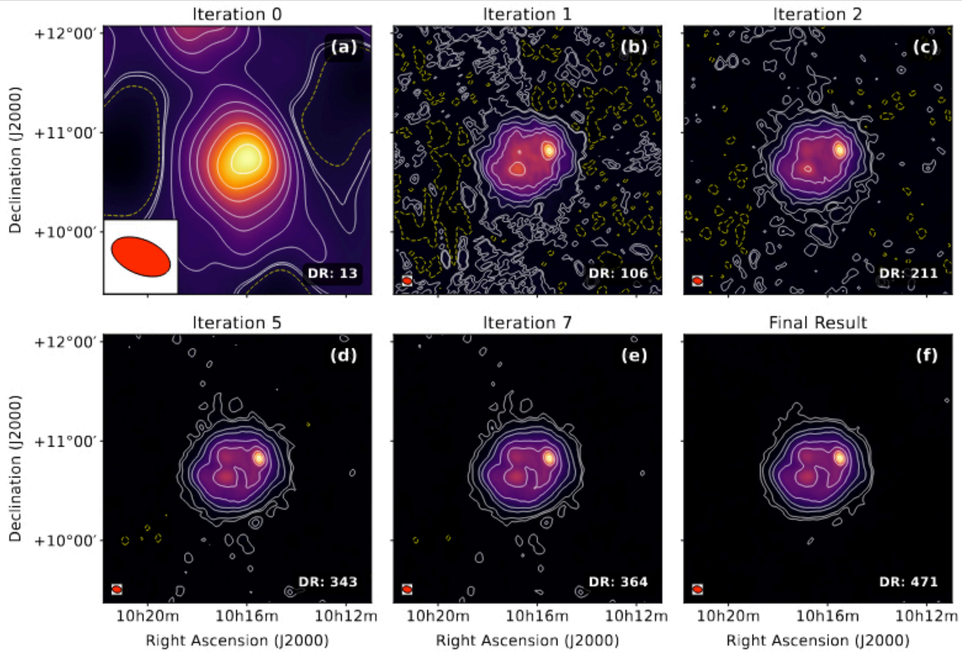
SIMPL:

- Runs in CWL and can return science ready data FITS HBA or LBA spectroscopic snapshot images to an archive.
- Fully automated and reliable (tested on a wide range of solar conditions and being tested with lots of archival data).
- Includes amplitude and phase calibration and optionally self calibration.
- Also includes refined flagging, calibrator selection and alignment of solar disk.

Ongoing development on polarisation calibration and using longer baselines.

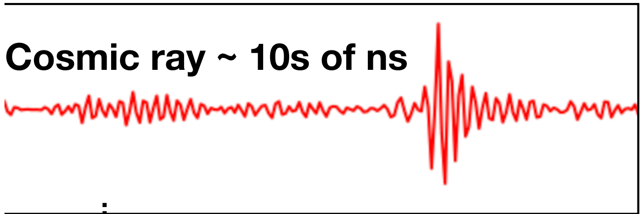
Slide adapted from P. Zucca:

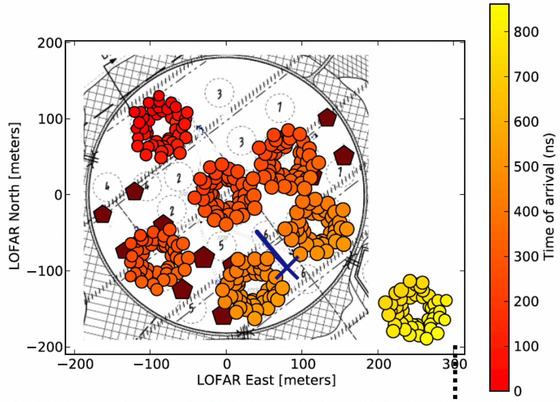


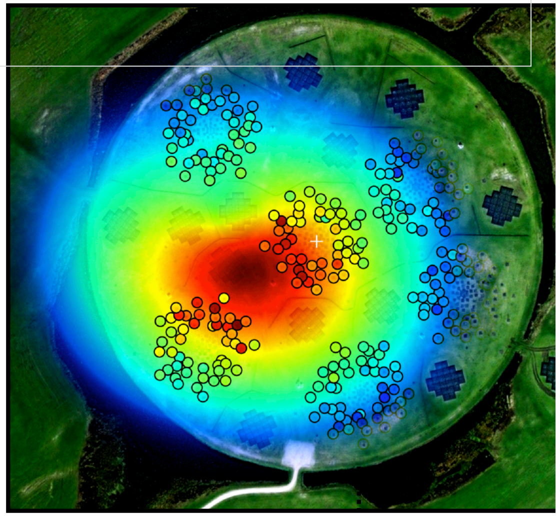


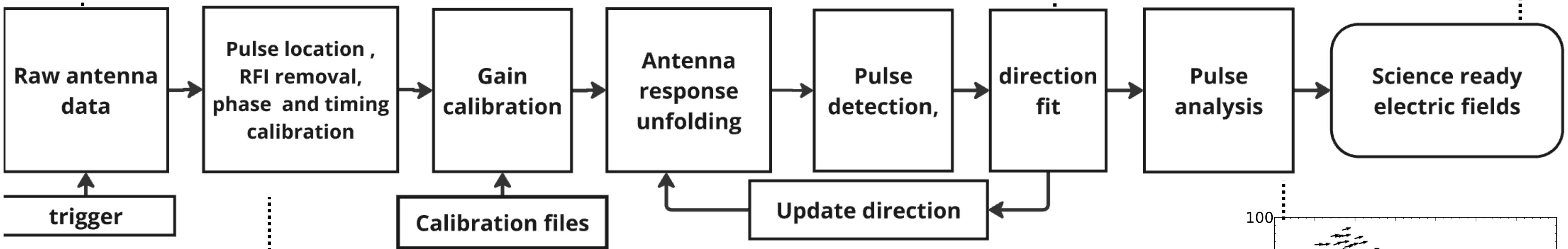
Cosmic Ray

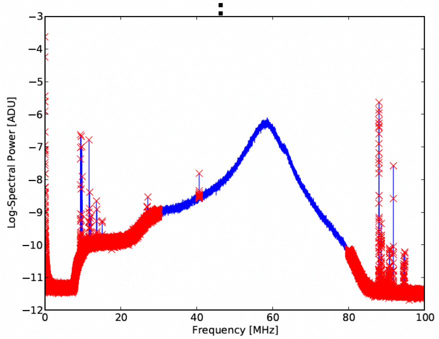
Pipeline builds upon the LOFAR1.0 pipeline (Schellart+ 2013) and characterises cosmic ray signals in the transient buffer data.

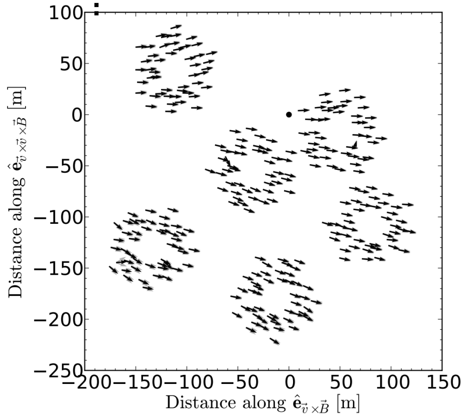












LOFAR 1.0 to LOFAR 2.0

- Pipeline integrated into widely used NuRadioReco software package
- Combined processing of HBA and LBA
- Lower energy events - extract weak signals

Slide adapted from K. Mulrey

LOFAR2.0 commissioning with pipelines

Main aim: demonstrate that LOFAR2.0 to as a system can meet its performance requirements

Only 9 LOFAR2.0 requirements make use of some of the pipelines that have been detailed (8 imaging and 1 Pulsar).

To pick a few of examples:

- … produce image cubes … with a rms noise level that is within 20% (50%) of the thermal noise… after 8hrs (1,000hrs) of observation
- Absolute flux calibration of the sources in the target field of a LOFAR observations shall be reproducible to within 2%
- LOFAR2.0shall have an imaging dynamic range of > 70 dB.
- LOFAR2.0 shall have an… imaging pipeline… generating calibrated image cubes… with a latency ≤3 days after completion of the observation.

LOFAR2.0 commissioning with pipelines

- Absolute flux calibration of the sources in the target field of a LOFAR observations shall be reproducible to within 2%

Test plan:

For this test we will conduct an observing run for each calibrator (3C196, 3C295, 3C147, 3C286, 3C48) consisting for 10mins calibrator at field centre, 10 mins at offsets corresponding to 60% and 30% of the power primary beam towards east and north. This entire procedure will be repeated three times in the same observing session. We will use this dataset for two tests:

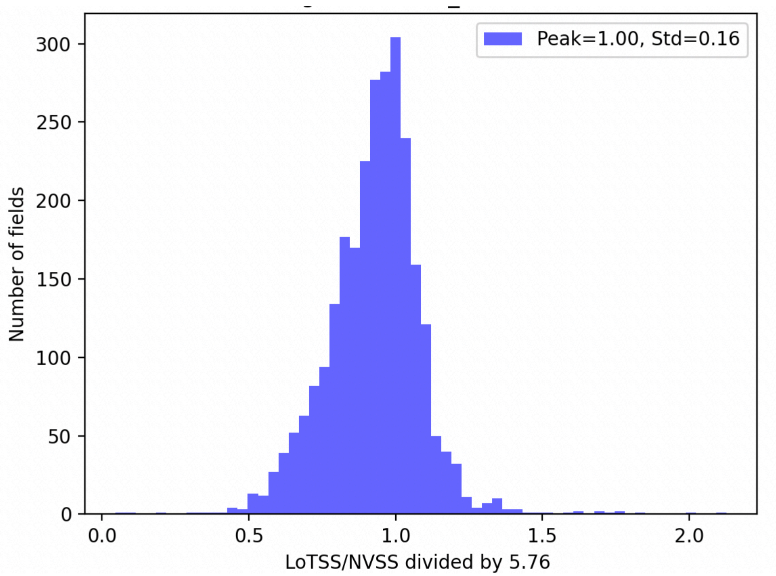
- For each central calibrator source we will calibrate the data in LINC-calibrator. We will then apply these calibrator solutions to the offset calibrator observings in the same observing run.
- For each central calibrator source we will calibrate the data in LINC-calibrator. We will then apply these calibrator solutions each of the other central calibrator sources which includes both different standard calibrators and the same calibrator but at a different sidereal time.

LOFAR2.0 commissioning with pipelines

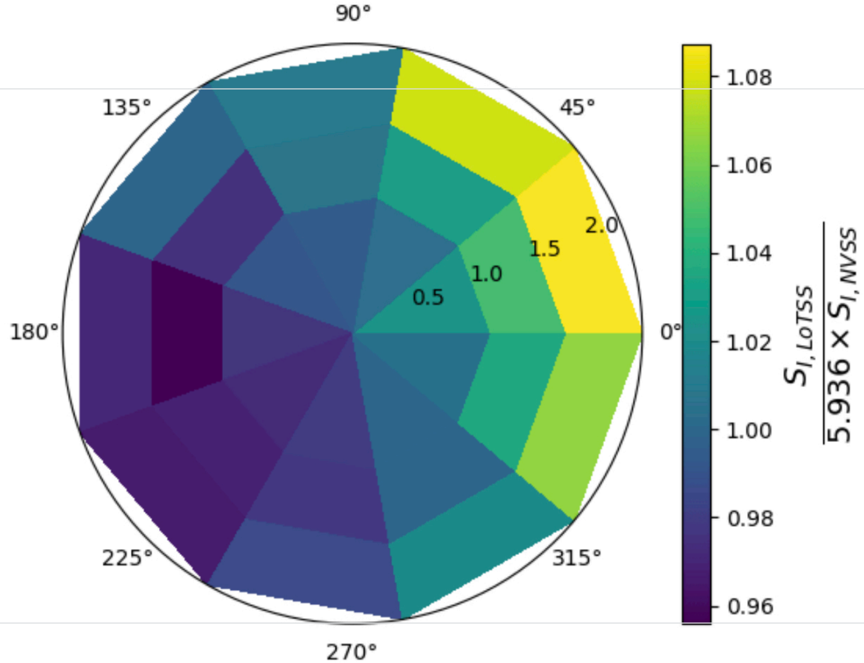
- Absolute flux calibration of the sources in the target field of a LOFAR observations shall be reproducible to within 2%.

Previous results:

Unable to reproduce flux density scale accurately. Finding that it varies across the field of view and that ~10% error remaining even after correcting.



Ratio of LoTSS (before much flux density scale tweaking) to NVSS integrated flux densities for 850 fields.



Ratio of LoTSS to NVSS integrated flux densities for 850 fields split into segments.

LOFAR2.0 commissioning with pipelines

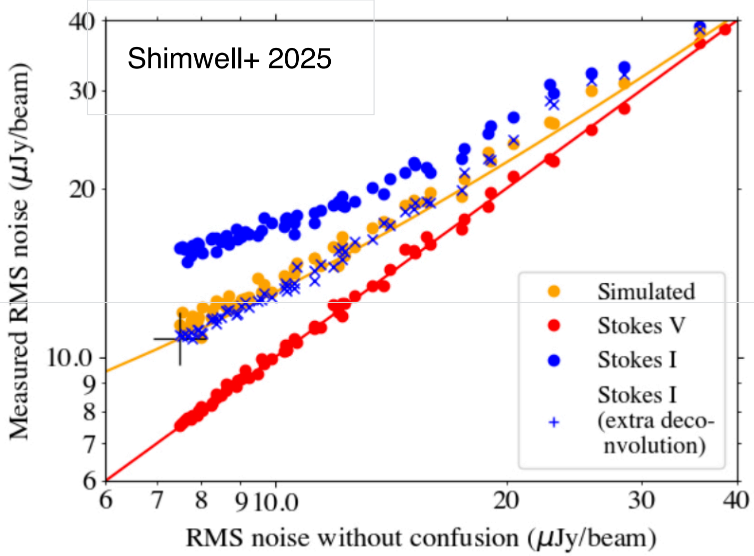
- … produce image cubes … with a rms noise level that is within 50% of the thermal noise… after 1,000hrs of observation

Draft test plan:

For HBA, we will run the data through LINC and DDF-pipeline and RAPTHOR, for LBA through LiLF, producing output I, Q, U, V data products at the specified spectral resolutions. We will use the Stokes V images to measure the instrumental noise (using an artifact free region outside of the main field of view) across the observed bandwidth. We will then compare the position dependent rms noise levels derived from the DDF-pipeline/RAPTHOR/LiLF I, Q, U, V data products to the single value per channel image noises derived from the artifact free region of the Stokes V cube. See how the rms noise decreases with observing time up to a reasonable limit (say 50 hrs)

Previous results:

Most data used in NL-baseline image ~550hrs. Thermal noise ~7.5uJy/beam and measured noise ~10.7uJy/beam with difference explainable by confusion noise. Unaware of any analysis on very deep image cubes.



Summary

Challenge of processing all LOFAR2.0 data to science quality in a timely manner is enormous. However, lots of optimism as:

- Excellent progress on all pipelines in preparation for LOFAR2.0.
- Tools and pipelines are continuous improving in efficiency and functionality

We are preparing to use these tools and pipelines to do some of the final stages of the commissioning once LOFAR-NL is ready for observing. Tests include seeing if we can reach performance requirements in dynamic range, image sensitivity across observing frequencies, flux density scale accuracy etc.

If you'd like to help conduct or plan the LOFAR2.0 commissioning tests that use pipelines email: shimwell@astron.nl or iacobelli@astron.nl