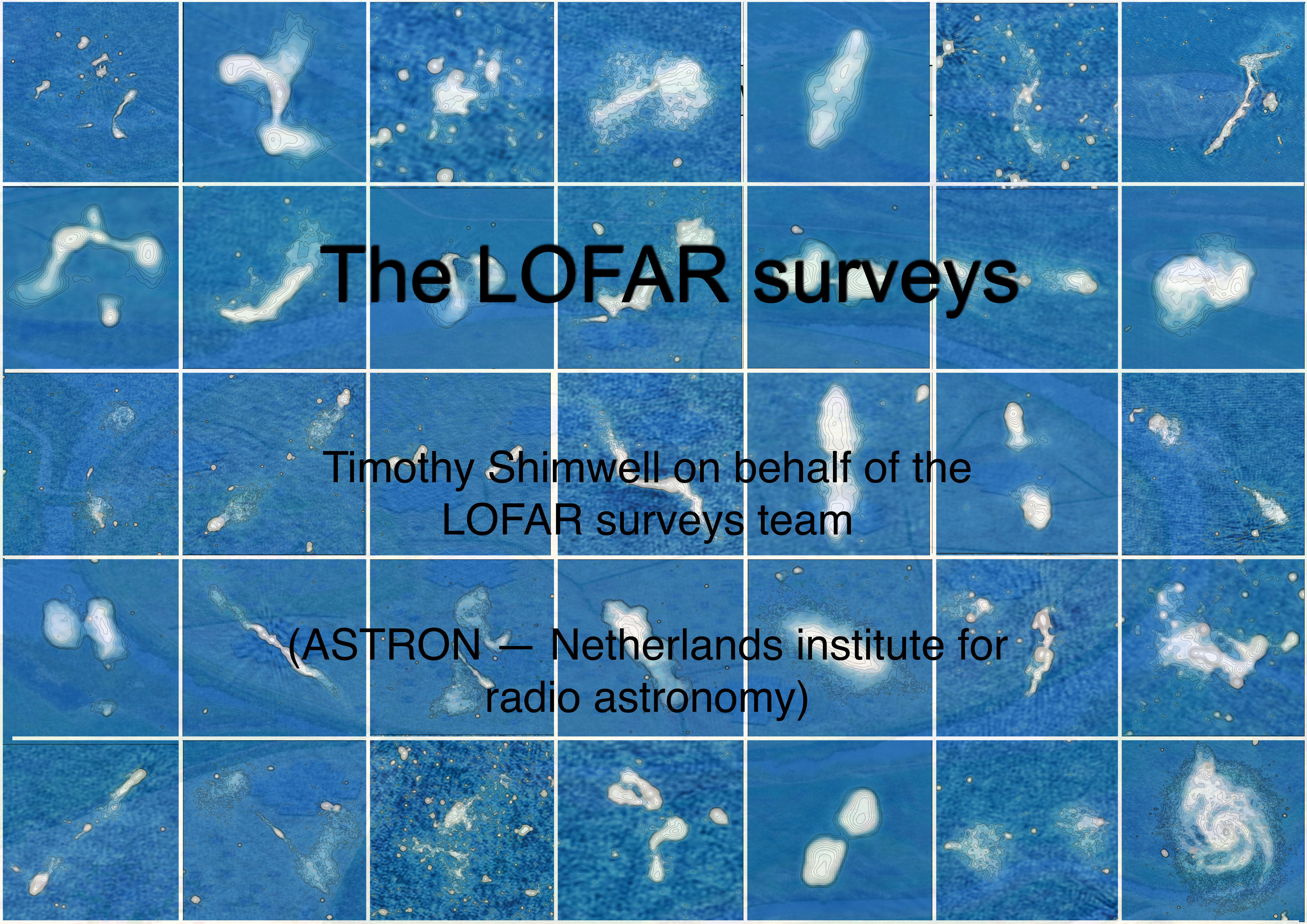


The LOFAR surveys

Timothy Shimwell on behalf of the
LOFAR surveys team

(ASTRON — Netherlands institute for
radio astronomy)

The LOFAR surveys team

190 members from ~30 institutes, over 6,000 hours of observations, well over 10PB of data and lots of hard work.

Scientific aims of the LOFAR surveys

Lead — Rottgering.

- Highest redshift radio sources: George Miley
- Clusters and cluster halo sources: Gianfranco Brunetti & Marcus Brüggen
- Evolution of AGN and star forming galaxies: Philip Best
- Detailed studies of low-redshift AGN: Raffaella Morganti
- Nearby Galaxies: Krzysztof Chyzy & John Conway
- Gravitational lensing: Neal Jackson
- Galactic radio sources: Glenn White & Marijke Haverkorn
- Cosmological studies: Matt Jarvis
- Other projects

Collaborations with the LOFAR magnetism, transient and the EoR groups.

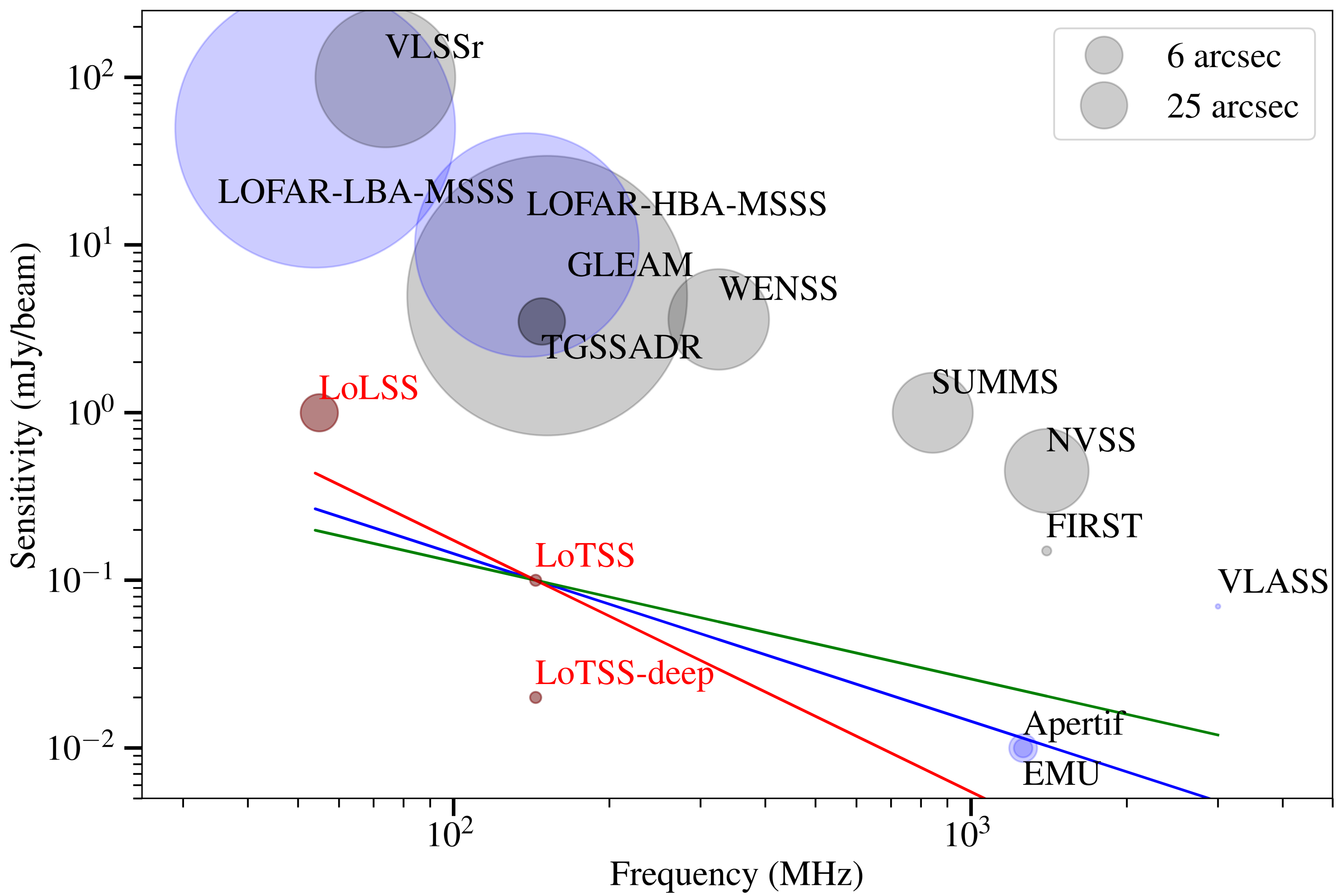
If you want to collaborate with the LOFAR surveys contact lotus-admin@strw.leidenuniv.nl

The LOFAR surveys

~14,000hrs of LOFAR HBA observations are required for LoTSS. 120-168MHz, 6" resolution, 0.1mJy/beam noise. Over 35% complete

Over 500hrs on LoTSS deep fields (primarily ELAIS-N1, Lockman hole and Bootes). 120-168MHz, 6" resolution, ~0.02mJy/beam noise

LBA survey, LoLSS, only preliminary region done to date (just a few percent). 42-66MHz, 15" resolution, 1mJy/beam noise) — de Gasperin talk.



Green $\alpha = -0.7$, blue $\alpha = -1.0$, red $\alpha = -1.5$

The LOFAR Two-metre Sky Survey (LoTSS) — data release 1

T. W. Shimwell, C. Tasse, M. J. Hardcastle, A. P. Mechev, W. L. Williams, P. N. Best, H. J. A. Rottgering, J. R. Callingham, T. J. Dijkema, F. de Gasperin, D. N. Hoang, B. Hugo, M. Mirmont, J. B. R. Oonk, I. Prandoni, D. Rafferty, J. Sabater, O. Smirnov, R. J. van Weeren, G. J. White, M. Atemkeng, L. Bester, E. Bonnassieux, M. Bruggen, G. Brunetti, K. T. Chyzy, R. Cochrane, J. E. Conway, J. H. Croston, A. Danezi, K. Duncan, M. Haverkorn, G. H. Heald, M. Iacobelli, H. T. Intema, N. Jackson, M. Jamrozy, M. J. Jarvis, R. Lakhoo, M. Mevius, G. K. Miley, L. Morabito, R. Morganti, D. Nisbet, E. Orru, S. Perkins, R. F. Pizzo, C. Schrijvers, D. J. B. Smith, R. Vermeulen, M. W. Wise, L. Alegre, D. J. Bacon, I. M. van Bemmel, R. J. Beswick, A. Bonafede, A. Botteon, S. Bourke, M. Brienza, G. Calistro Rivera, R. Cassano, A. O. Clarke, C. J. Conselice, R. J. Dettmar, A. Drabent, C. Dumba, K. L. Emig, T. A. Ensslin, C. Ferrari, M. A. Garrett, R. T. Genova-Santos, A. Goyal, G. Gurkan, C. Hale, J. J. Harwood, V. Heesen, M. Hoeft, C. Horellou, C. Jackson, G. Kokotanekov, R. Kondapally, M. Kunert-Bajraszewska, V. Mahatma, E. K. Mahony, S. Mandal, J. P. McKean, A. Merloni, B. Mingo, S. Mooney, B. Nikiel-Wroczynski, S. P. O'Sullivan, J. Quinn, W. Reich, C. Roskowinski, A. Rowlinson, F. Savini, A. Saxena, D. J. Schwarz, A. Shulevski, S. S. Sridhar, H. R. Stacey, S. Urquhart, M. H. D. van der Wiel, E. Varenius, B. Webster, A. Wilber

W. L. Williams, M. J. Hardcastle, P. N. Best, J. Sabater, J. H. Croston, K. J. Duncan, T. W. Shimwell, H. J. A. Rottgering, D. Nisbet, G. Gurkan, L. Alegre, R. K. Cochrane, A. Goyal,C. L. Hale, N. Jackson, M. Jamrozy, R. Kondapally, M. Kunert-Bajraszewska, V. H. Mahatma, B. Mingo, L. K. Morabito, I. Prandoni, C. Roskowinski, A. Shulevski, D. J. B. Smith, C. Tasse, S. Urquhart, B. Webster, G. J. White, R. J. Beswick, J. R. Callingham, K. T. Chyzy, F. de Gasperin, J. J. Harwood, M. Hoeft, M. Iacobelli, J. P. McKean, A. P. Mechev, G. K. Miley, D. J. Schwarz, R. J. van Weeren

Kenneth J Duncan, J. Sabater, H. J. A. Rottgering, M. J. Jarvis, D. J. B. Smith, P. N. Best, J. R. Callingham, R. Cochrane, J. H. Croston, M. J. Hardcastle, B. Mingo, L. Morabito, D. Nisbet, I. Prandoni, T. W. Shimwell, C. Tasse, G. J. White, W. L. Williams, L. Alegre, K. T. Chyzy, G. Gurkan, M. Hoeft, R. Kondapally, A. P. Mechev, G. K. Miley, D. J. Schwarz, R. J. van Weeren

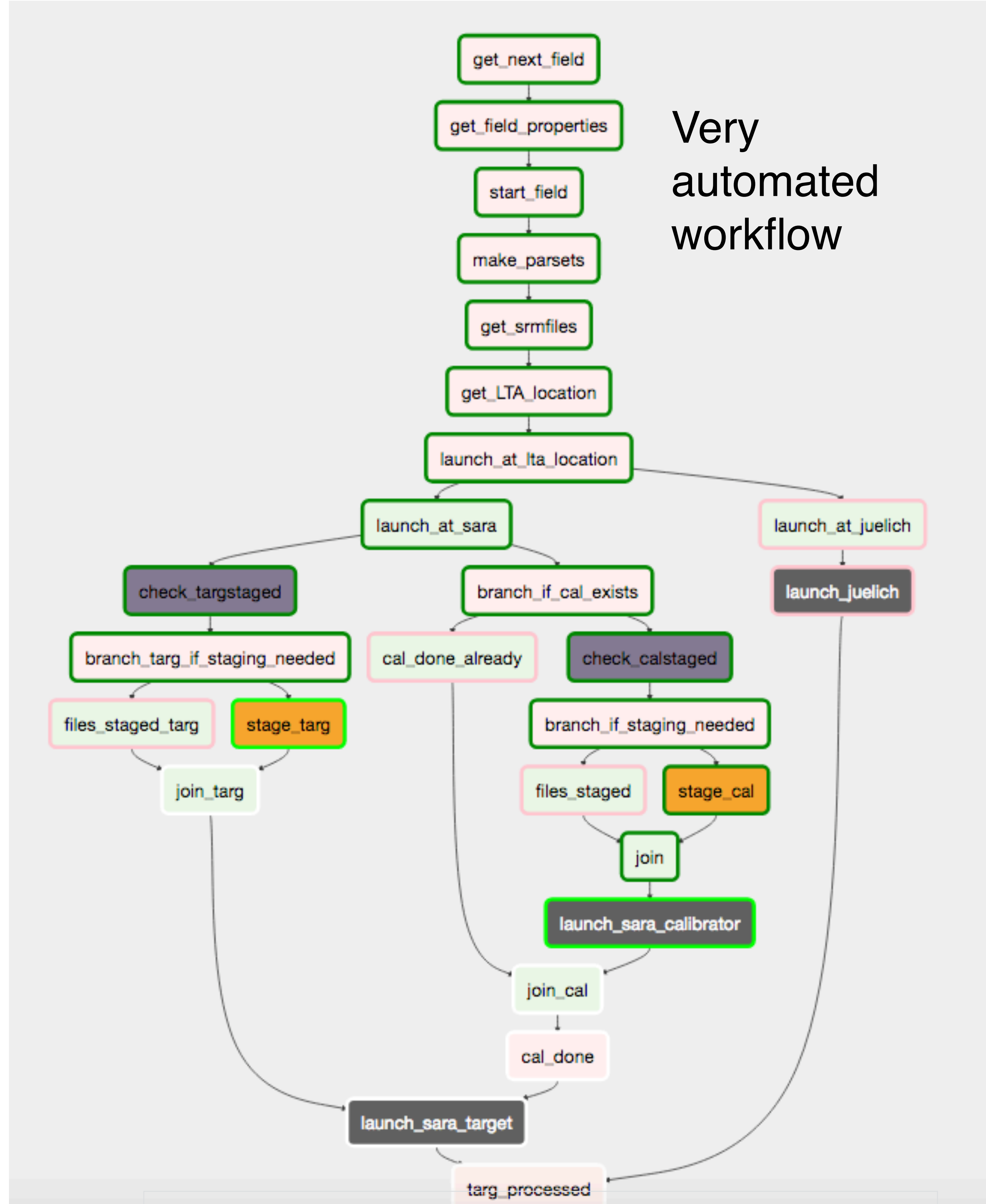
The data processing pipelines used for LoTSS

All data processing pipelines for LoTSS have been made public.

To process the >10PB of existing data the pipeline is implemented on the LOFAR archive compute facilities. This allows us to process ~16Tb of data within ~6hrs on each site.

Around 50% of existing LoTSS data have been processed with this pipeline.

The main bottleneck in data processing is getting the data out of the LOFAR archive.

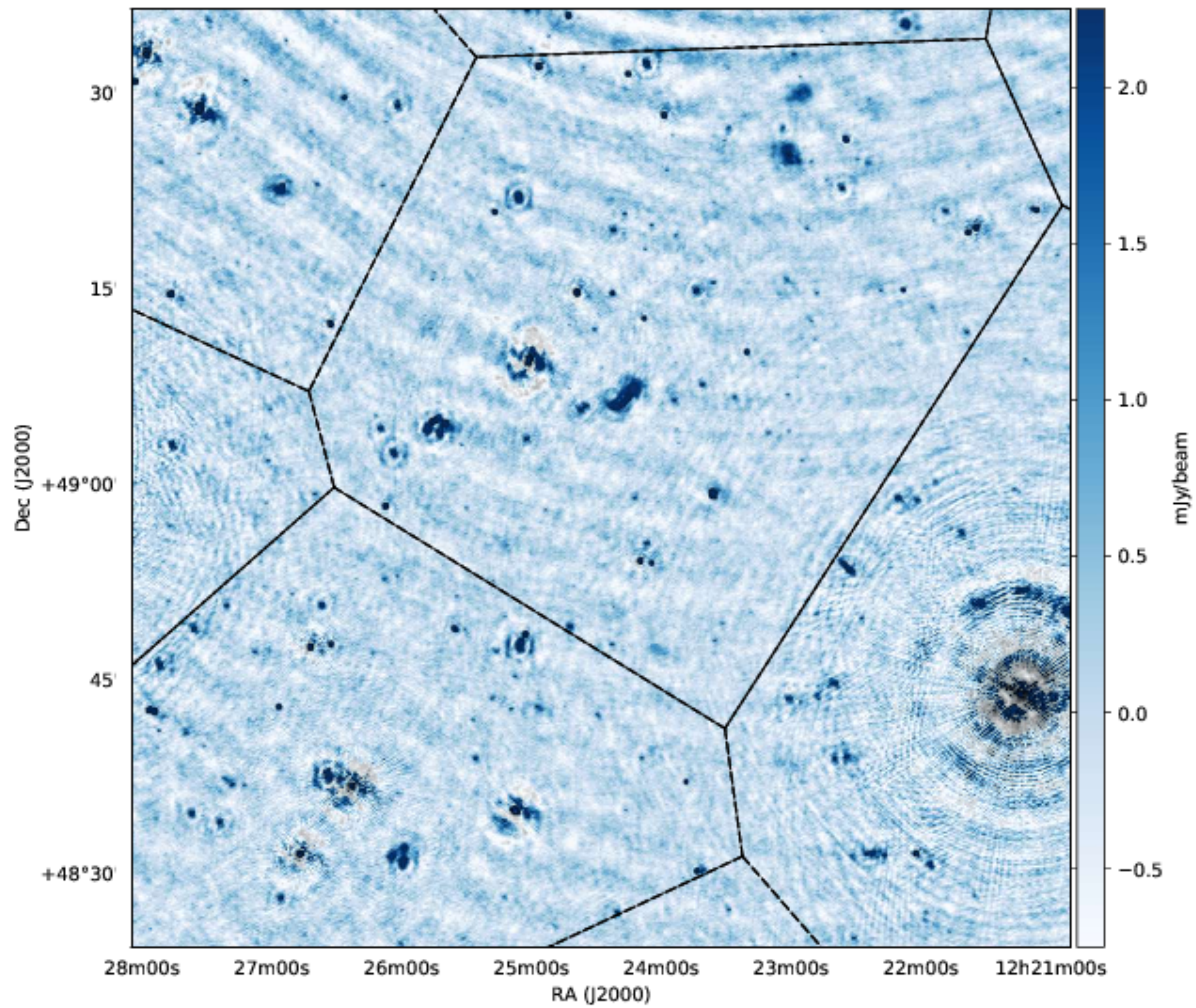


Very automated workflow

LTA processing: Mechev, Drabent, Oonk, Shimwell, et al.

The data processing pipelines used for LoTSS

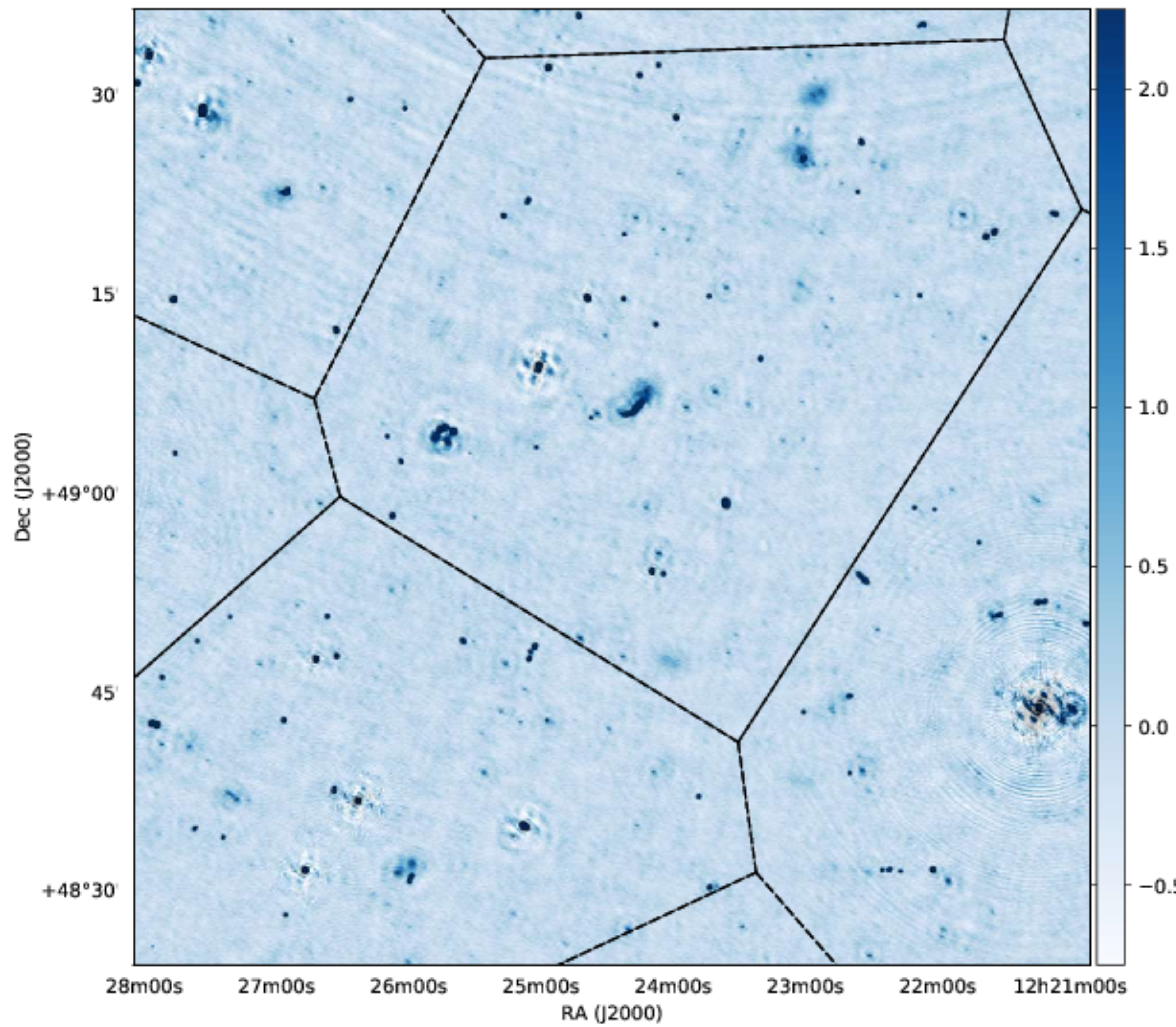
- .Wide field image created covering the full field of view
- Image tessellated to define facets.
- Calibration solutions obtained for all directions simultaneously using the model from entire wide field image.
- Imaging repeated with direction dependent calibration solutions applied and a better model constructed.
- Several self calibration cycles performed.



DDF-pipeline is public. It makes use of DDFacet and kMS for calibration and imaging (Tasse+ 2014, Smirnov+ 2015, and Tasse+ 2017). It produces science quality images in 4-6 days of processing but requires powerful (>200GB) nodes.

The data processing pipelines used for LoTSS

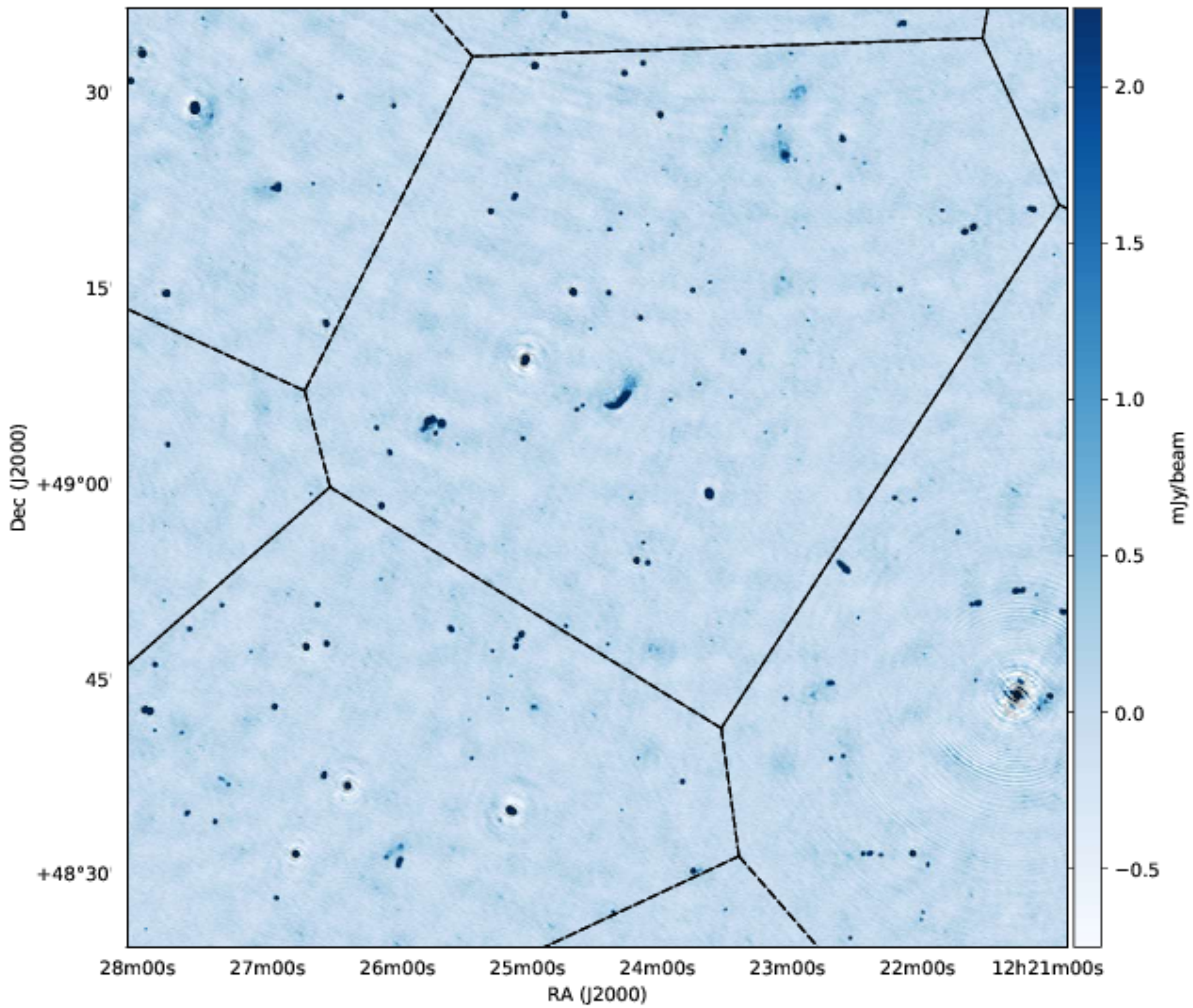
- .Wide field image created covering the full field of view
- Image tessellated to define facets.
- Calibration solutions obtained for all directions simultaneously using the model from entire wide field image.
- Imaging repeated with direction dependent calibration solutions applied and a better model constructed.
- Several self calibration cycles performed.



DDF-pipeline is public. It makes use of DDFacet and kMS for calibration and imaging (Tasse+ 2014, Smirnov+ 2015, and Tasse+ 2017). It produces science quality images in 4-6 days of processing but requires powerful (>200GB) nodes.

The data processing pipelines used for LoTSS

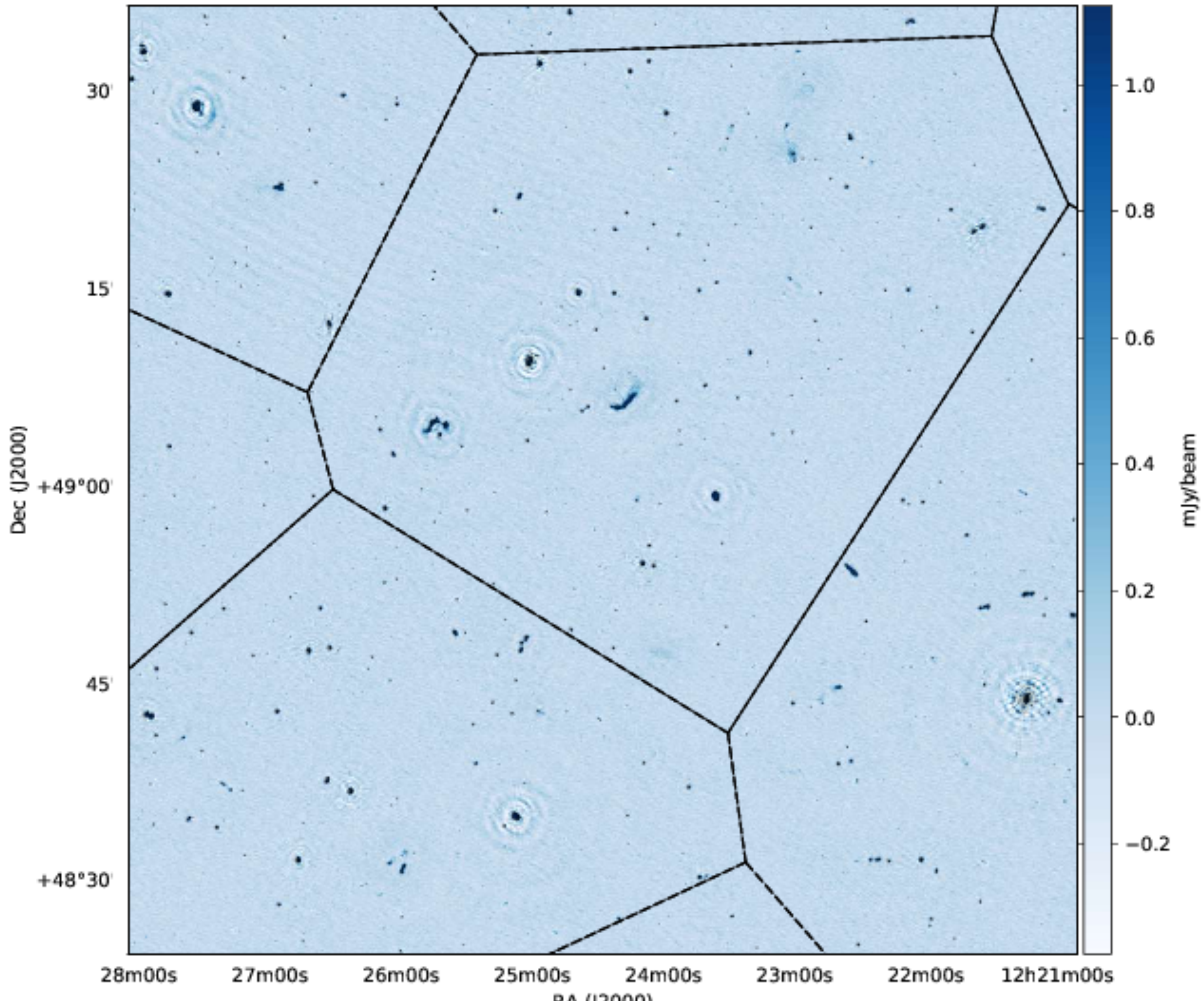
- .Wide field image created covering the full field of view
- Image tessellated to define facets.
- Calibration solutions obtained for all directions simultaneously using the model from entire wide field image.
- Imaging repeated with direction dependent calibration solutions applied and a better model constructed.
- Several self calibration cycles performed.



DDF-pipeline is public. It makes use of DDFacet and kMS for calibration and imaging (Tasse+ 2014, Smirnov+ 2015, and Tasse+ 2017). It produces science quality images in 4-6 days of processing but requires powerful (>200GB) nodes.

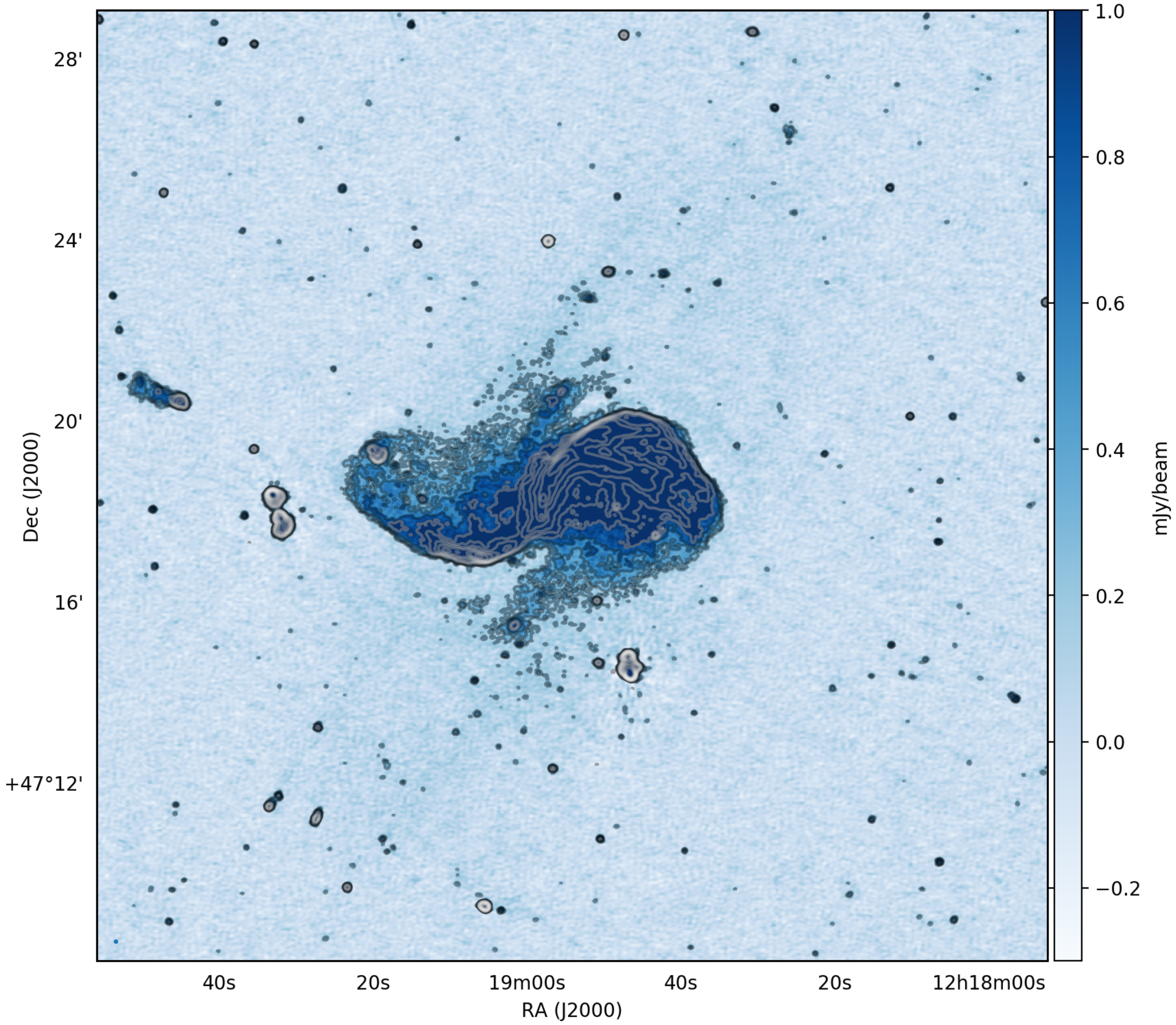
The data processing pipelines used for LoTSS

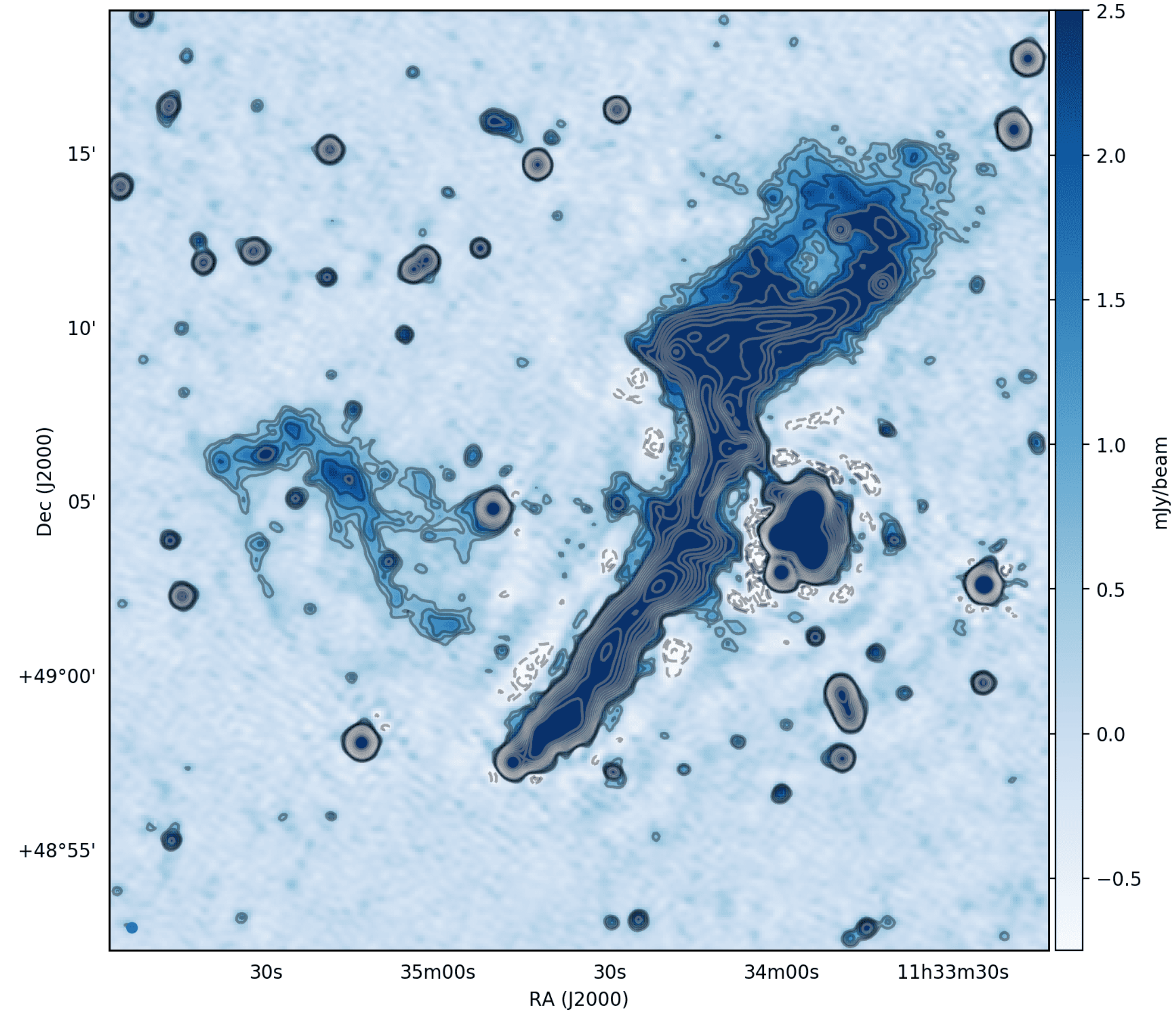
- .Wide field image created covering the full field of view
- Image tessellated to define facets.
- Calibration solutions obtained for all directions simultaneously using the model from entire wide field image.
- Imaging repeated with direction dependent calibration solutions applied and a better model constructed.
- Several self calibration cycles performed.

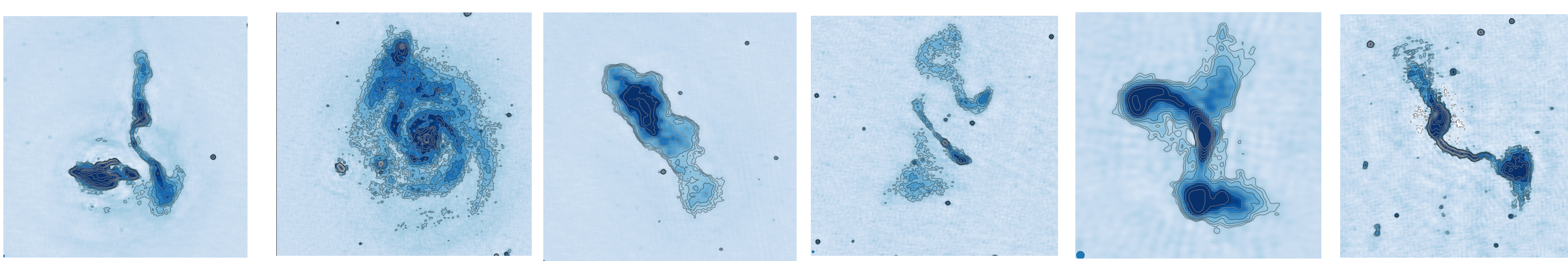


DDF-pipeline is public. It makes use of DDFacet and kMS for calibration and imaging (Tasse+ 2014, Smirnov+ 2015, and Tasse+ 2017). It produces science quality images in 4-6 days of processing but requires powerful (>200GB) nodes.

The data processing pipelines used for LoTSS

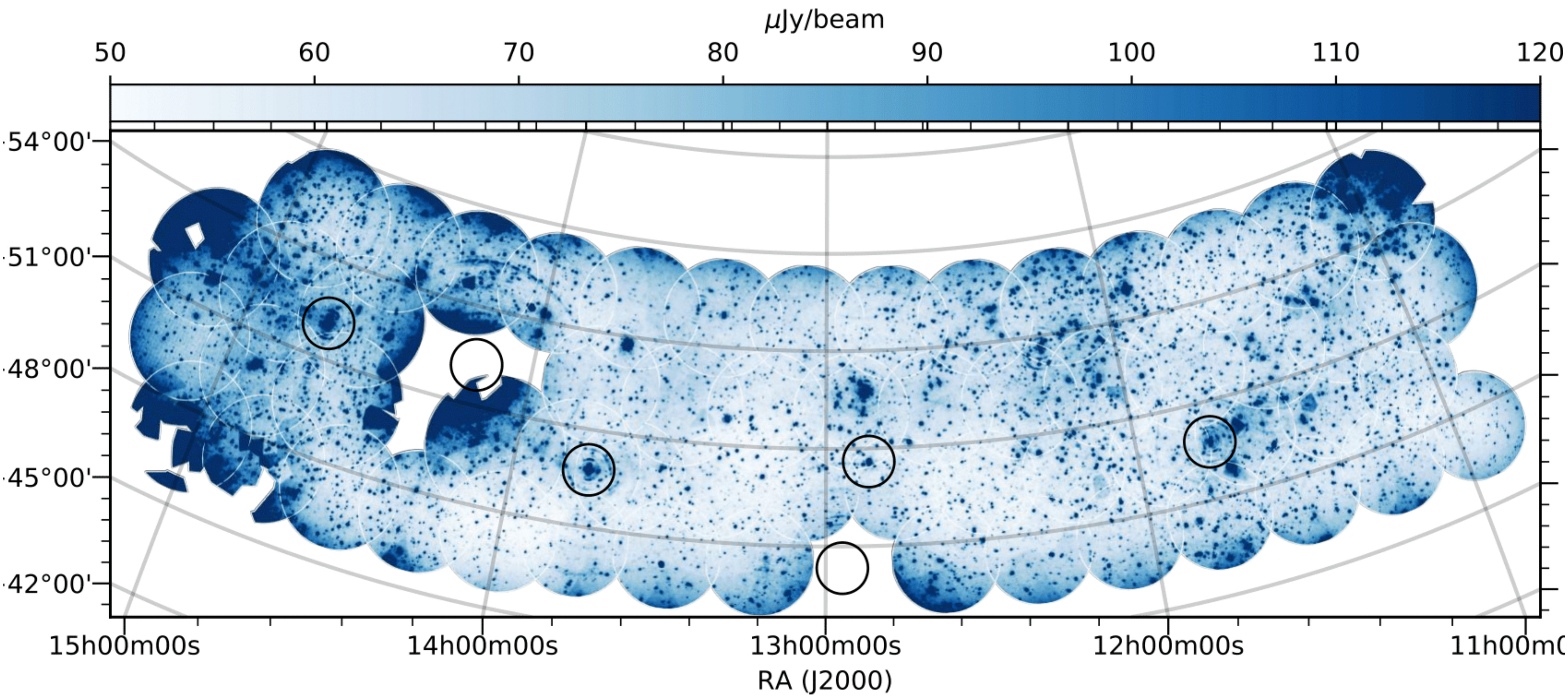






Example image after DD calibration. Note that improving direction dependent calibration and imaging routines is ongoing research.

LoTSS data release 1 — visit lofar-surveys.org



6" resolution, a median noise of 0.07mJy/beam, a source density 10 times higher than NVSS or FIRST, 90% complete at 0.45mJy/beam and astrometric accuracy of 0.2".

LoTSS-DR1 was published Feb 2019 together with ~25 science papers from the SKSP. The catalogues, images and all data processing software and pipelines are fully public.

LoTSS data release 1 — visit lofar-surveys.org

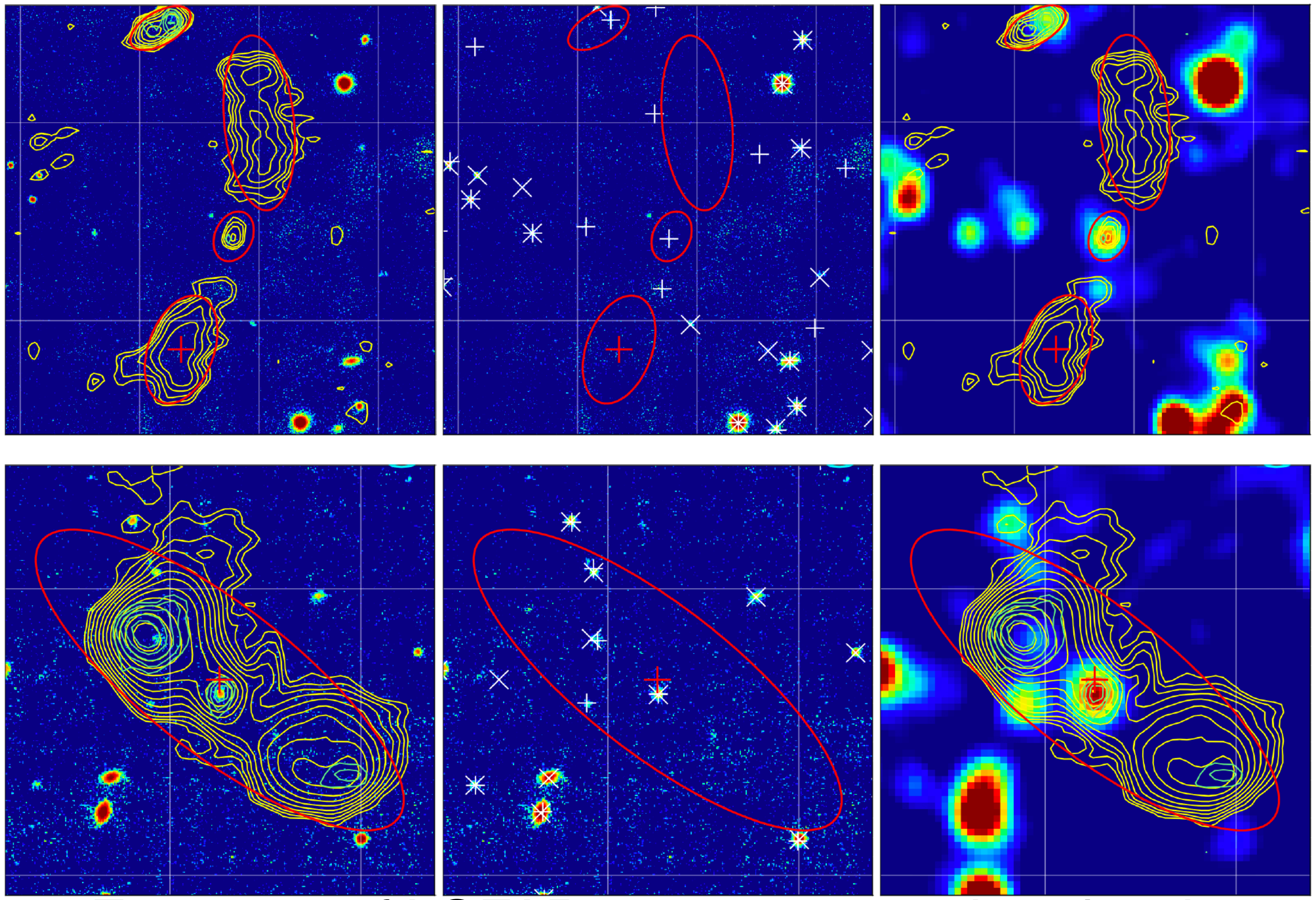
325,694 entires in the raw PyBDSF catalogue.

Corresponds to 318,520 radio sources after deblending, artefact rejection and joining multiple component sources (including extensive efforts to visually inspect ~10,000 sources) — see Williams talk

231,716 have counterparts in Pan-STARRS or WISE and for these photometric redshifts are are estimated.

	Number	Number with ID	ID fraction
All Sources	318,520	231,716	0.73
LR	299,730	221,269	0.74
LGZ	11,989	7,144	0.60
Deblending	2,435	2,338	0.96
Bright galaxy	965	965	1.00
No ID possible	3,401	0	0.00

The final LoTSS-DR1 catalogue contains radio sources, optical counter parts and photometric redshifts.



Examples of LOFAR galaxy zoo entries showing LoTSS, FIRST, Pan-STARRS and WISE images.

Williams, Duncan, Hardcastle, Sabatar et al.

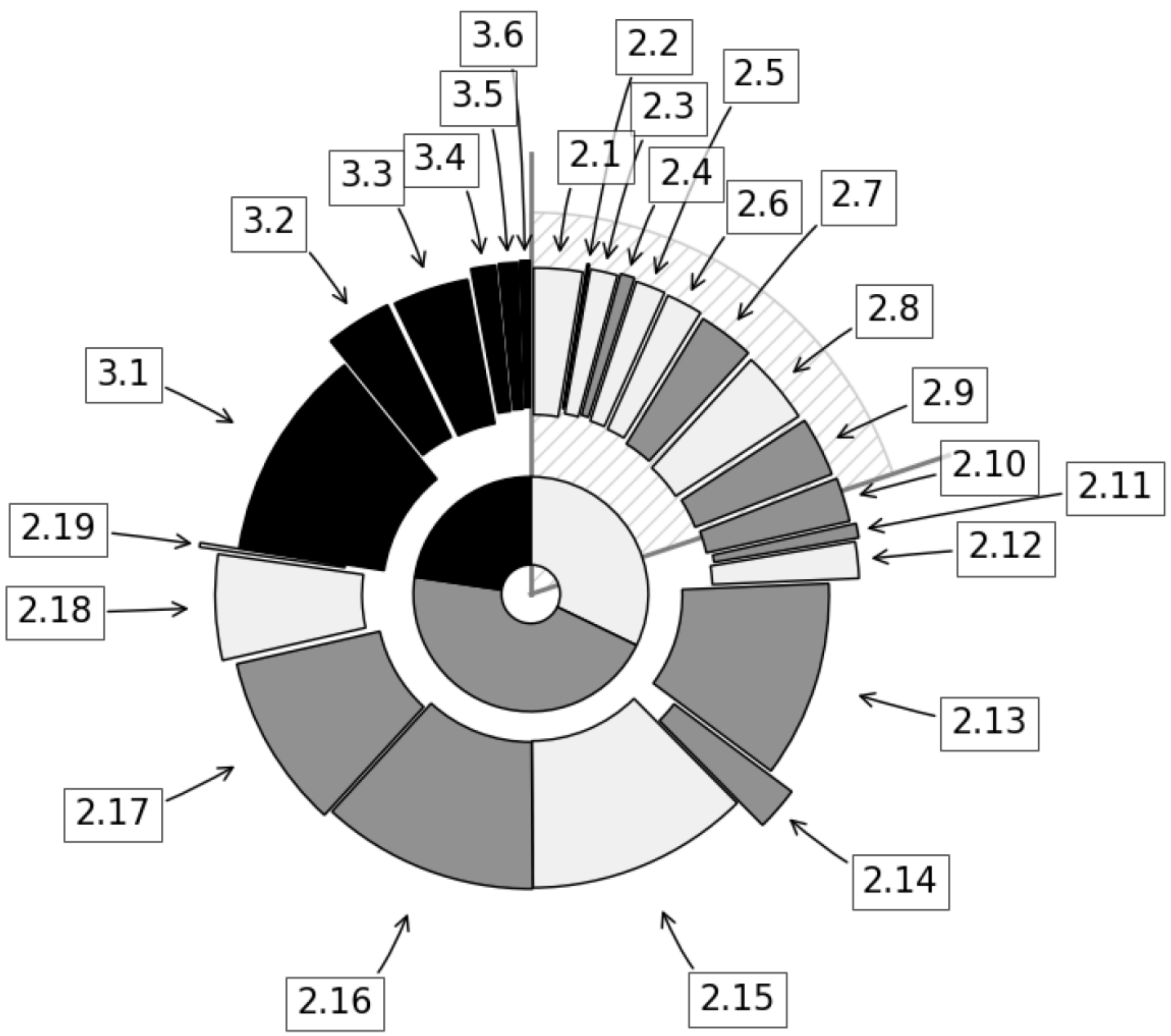
LoTSS data release 1

All LoTSS-DR1 products (images, raw radio catalogues, cross matched source catalogues, redshift estimates) are public. Release is described in Shimwell+ 2019, Williams+ 2019 and Duncan+ 2019 and was accompanied by ~25 scientific papers in a special issue of A&A and public release of all codes.

A new pipeline for LoTSS-DR2

Nearly a year of development went the new DR2 pipeline.

Approximately 14 nodes are running the pipeline in parallel to produce more than 2 fully calibrated images each day (90% of the processing is done in Hertfordshire, the remaining 10% in Leiden, Bologna and Hamburg)



Steps in the LoTSS pipeline by runtime (6 days in total on a good node). Light and dark areas are imaging and calibration steps respectively. Black is miscellaneous (Tasse+ in prep)

Main differences:

Deeper deconvolution — More very faint sources (about 10% more)

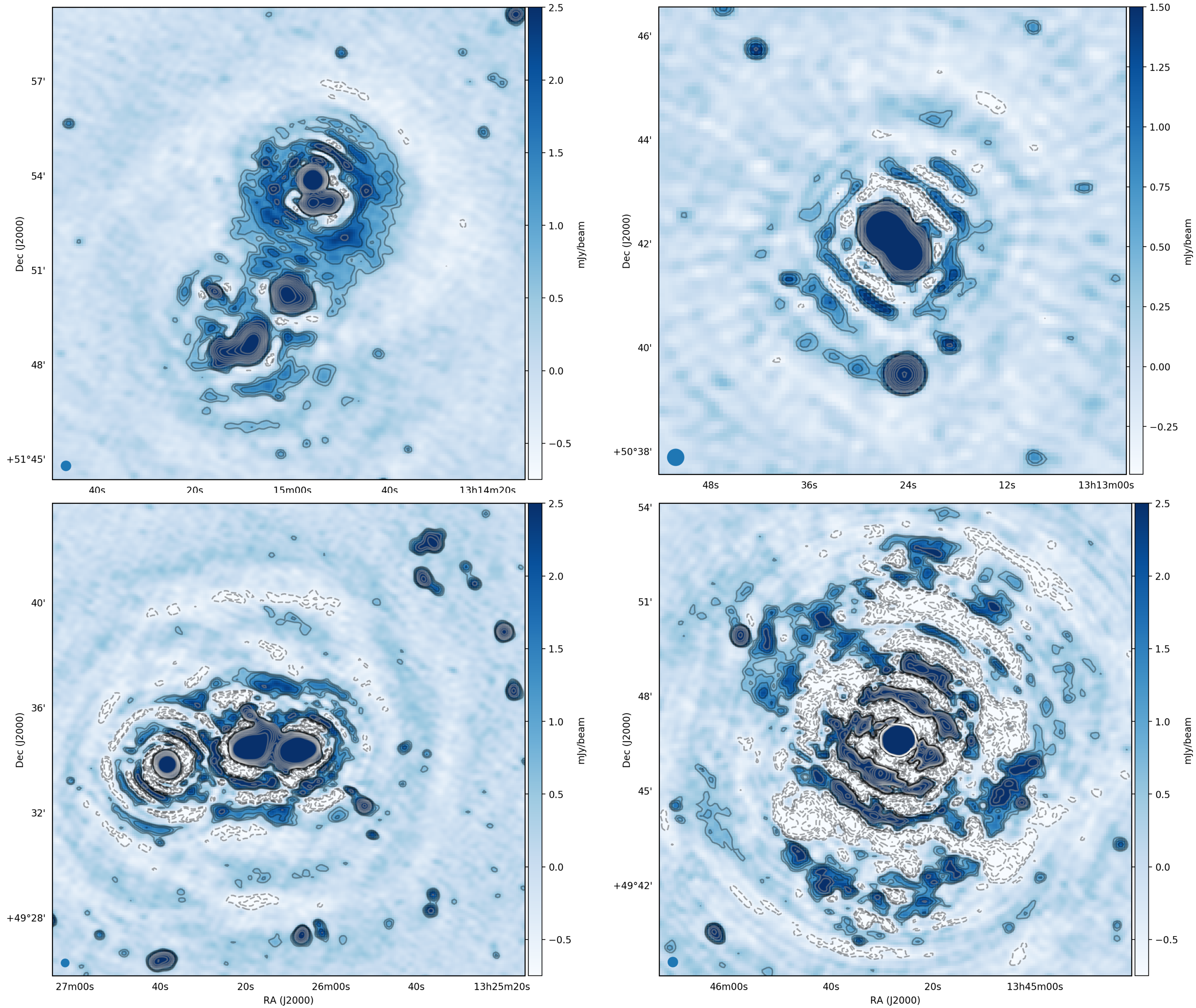
Redoing the direction independent calibration during direction dependent pipeline — Higher dynamic range

Better deconvolution and shorter uv-min in the calibration — No more artificial halos and holes

DDF-pipeline: Tasse, Hardcastle, Shimwell

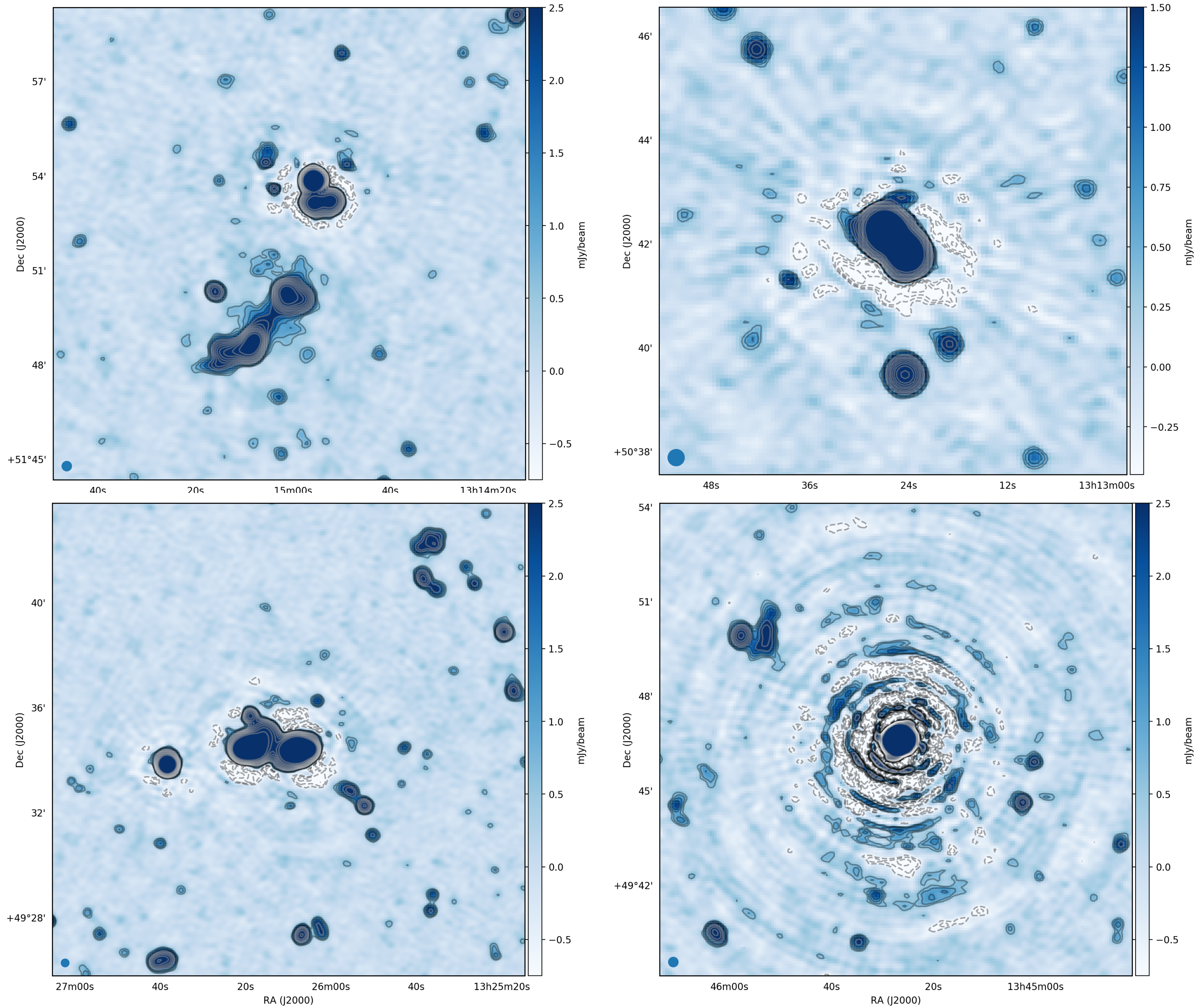
Improvements in LoTSS-DR2

DR1 has quite a low dynamic range (~10% of the area has a >15% enhanced noise) which resulted in an ~5% failure rate for extragalactic fields



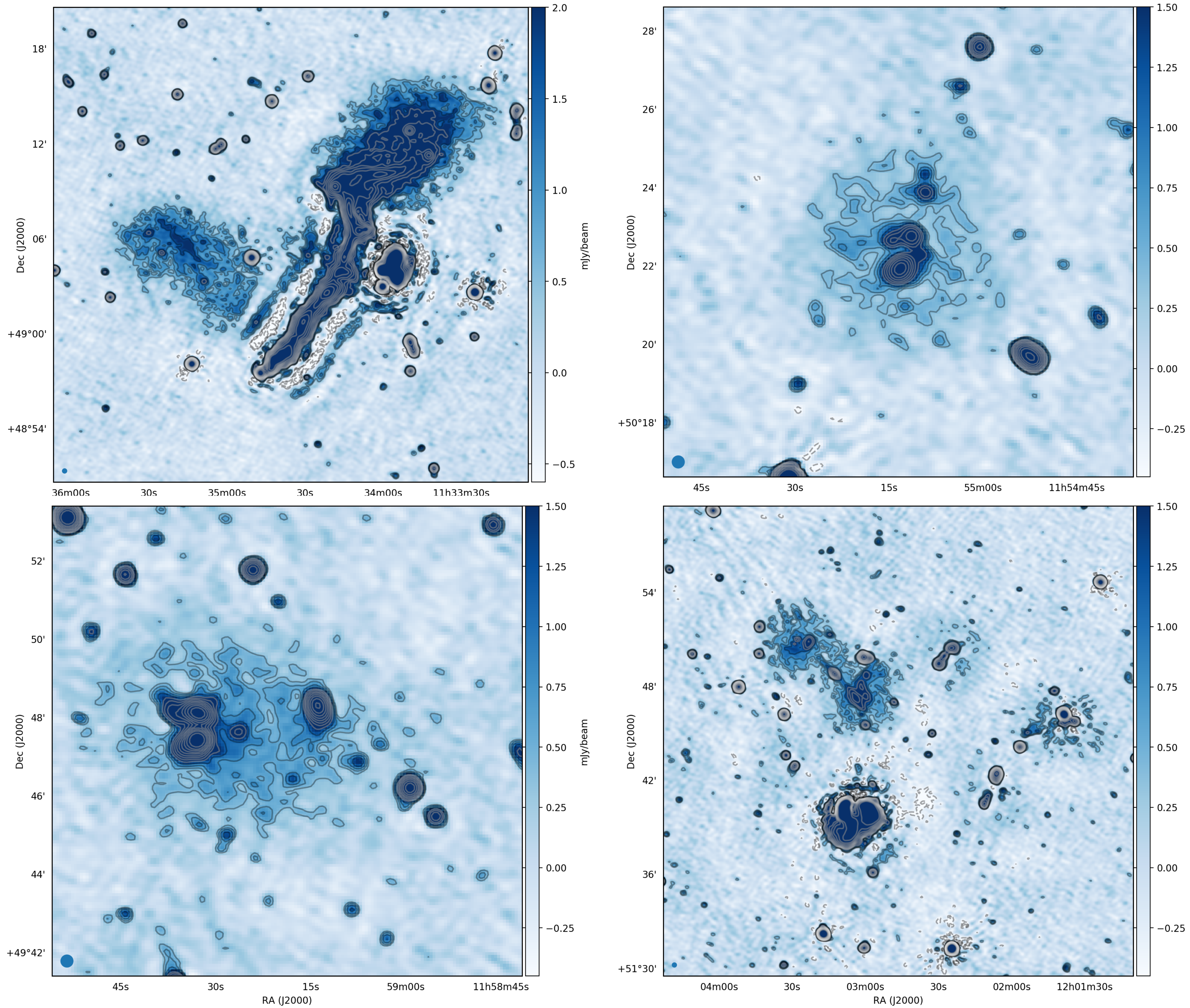
Improvements in LoTSS-DR2

DR2 improves the dynamic range significantly.



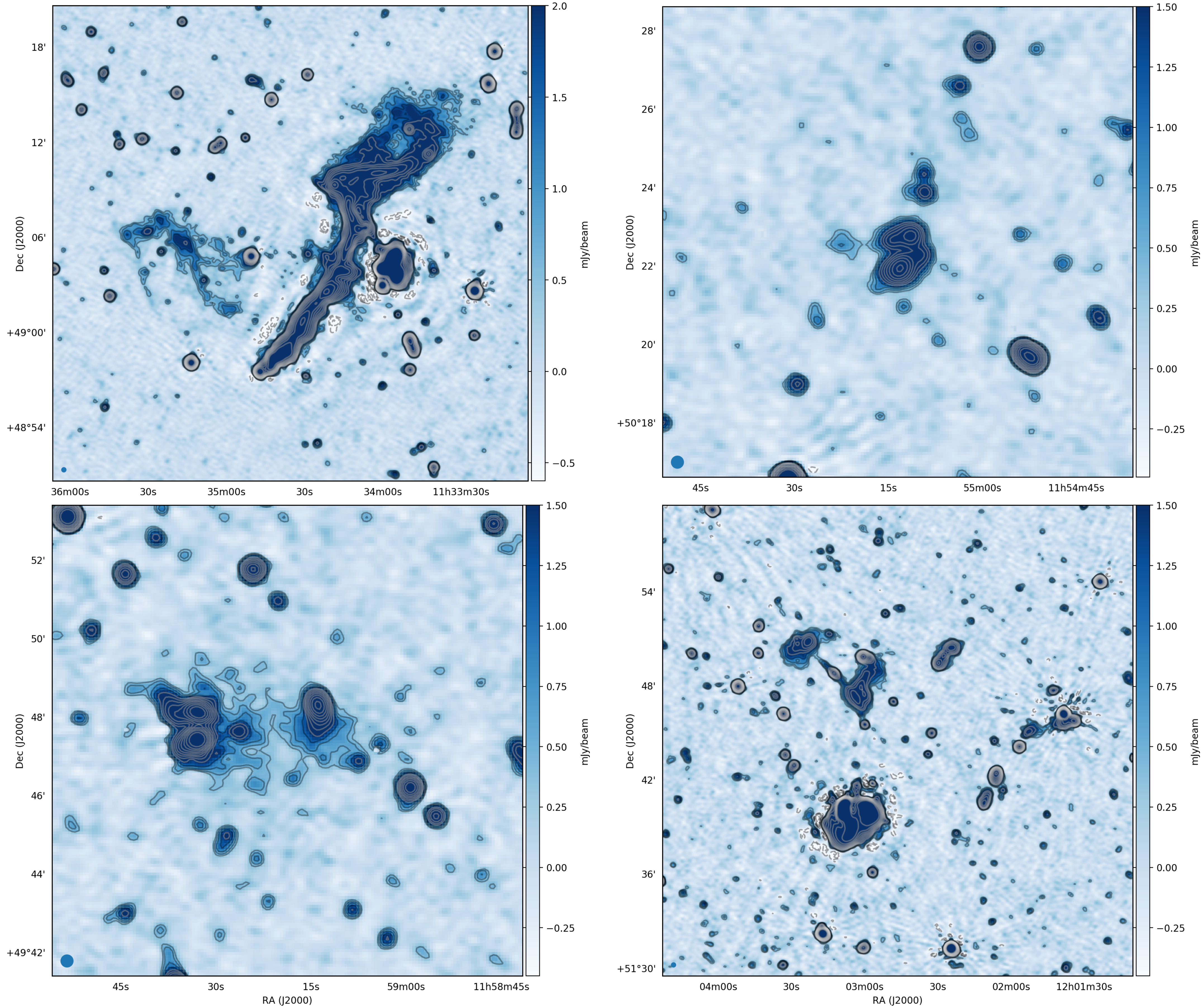
Improvements in LoTSS-DR2

DR1 has artificial halos and holes mainly around extended sources.

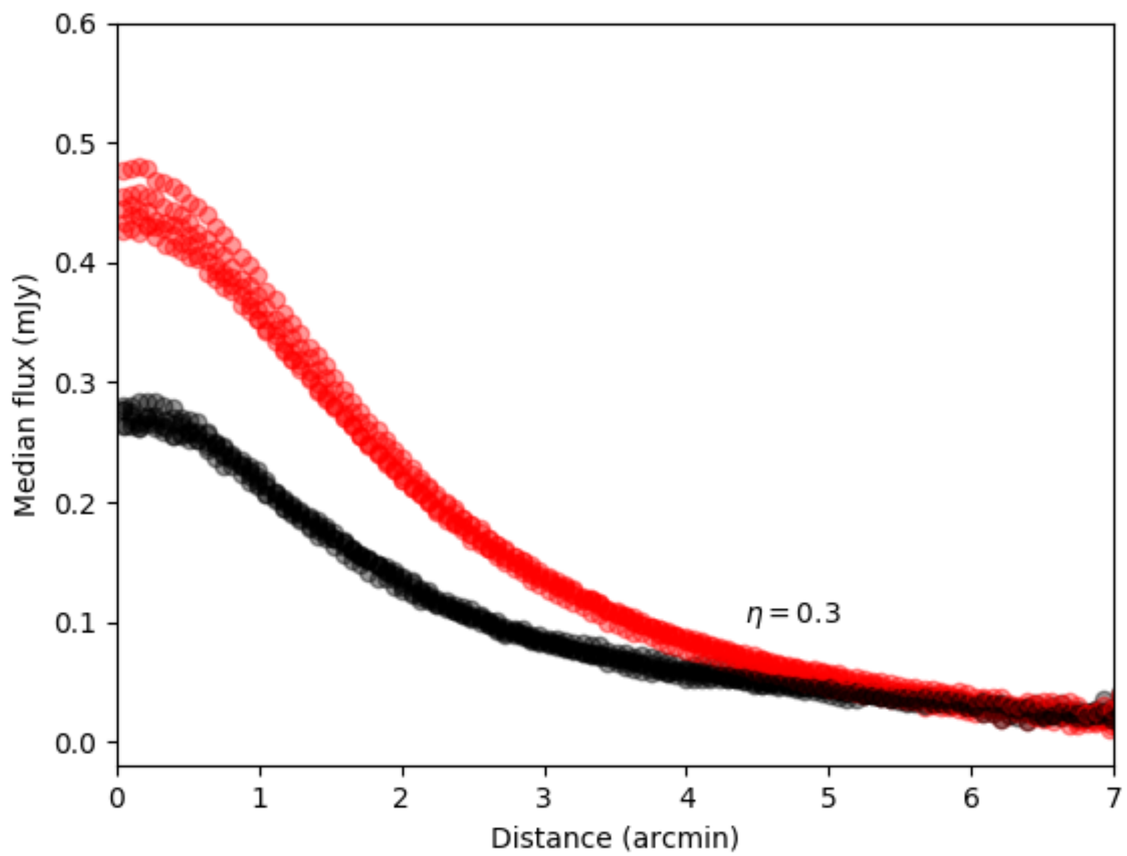


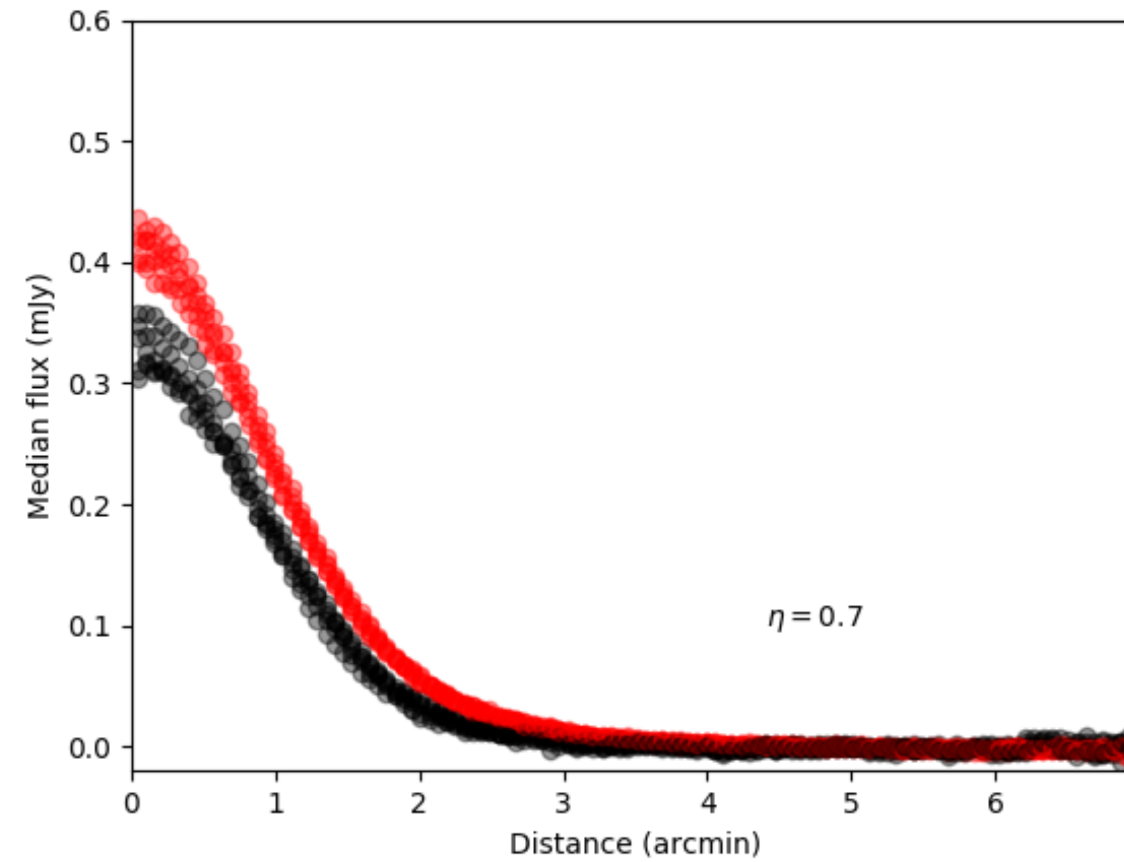
Improvements in LoTSS-DR2

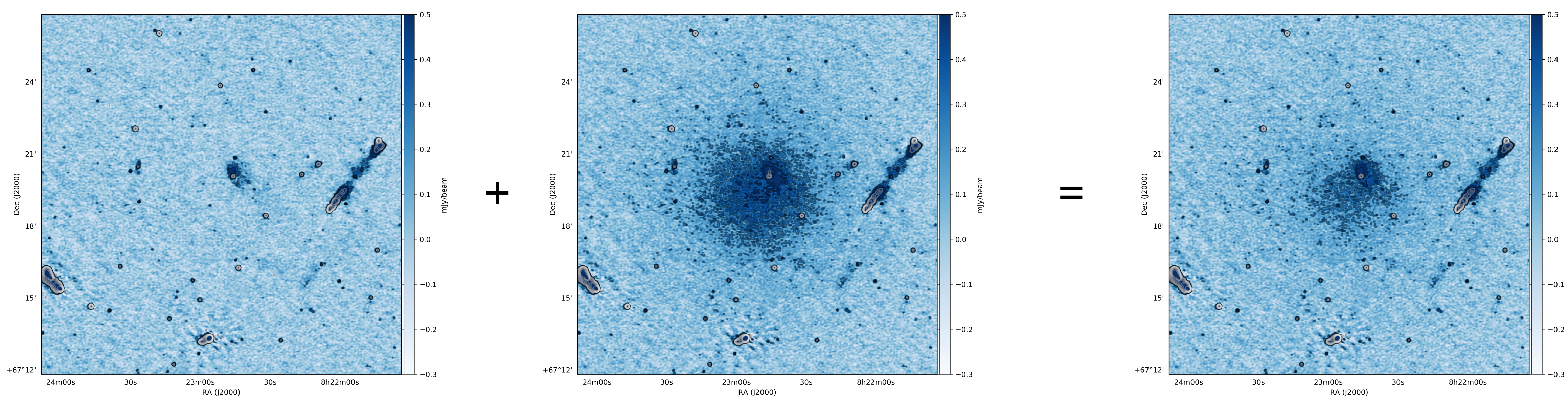
DR2 essentially removes these artificial halos and holes.



Improved fidelity and recovery of unmodelled diffuse emission

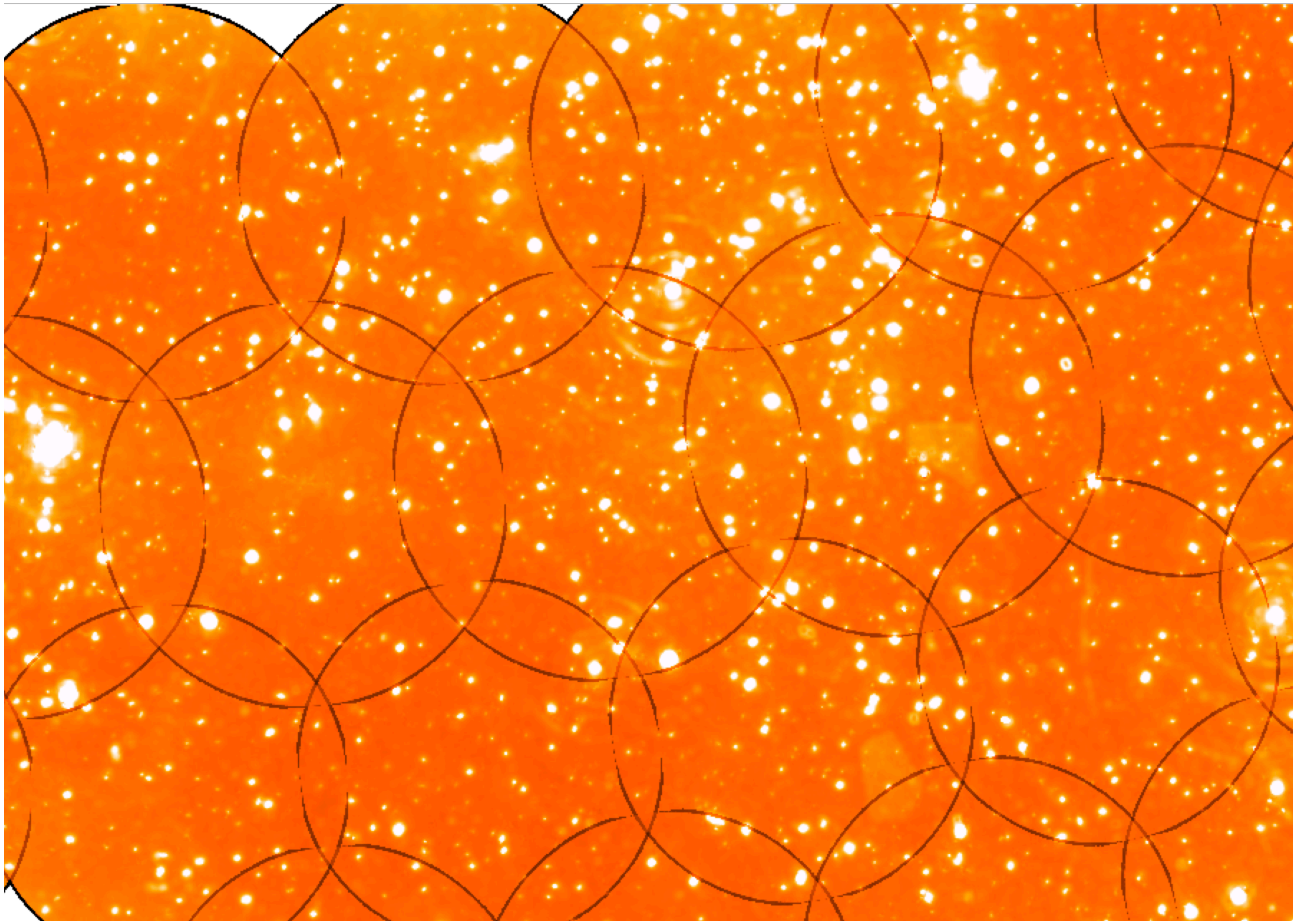






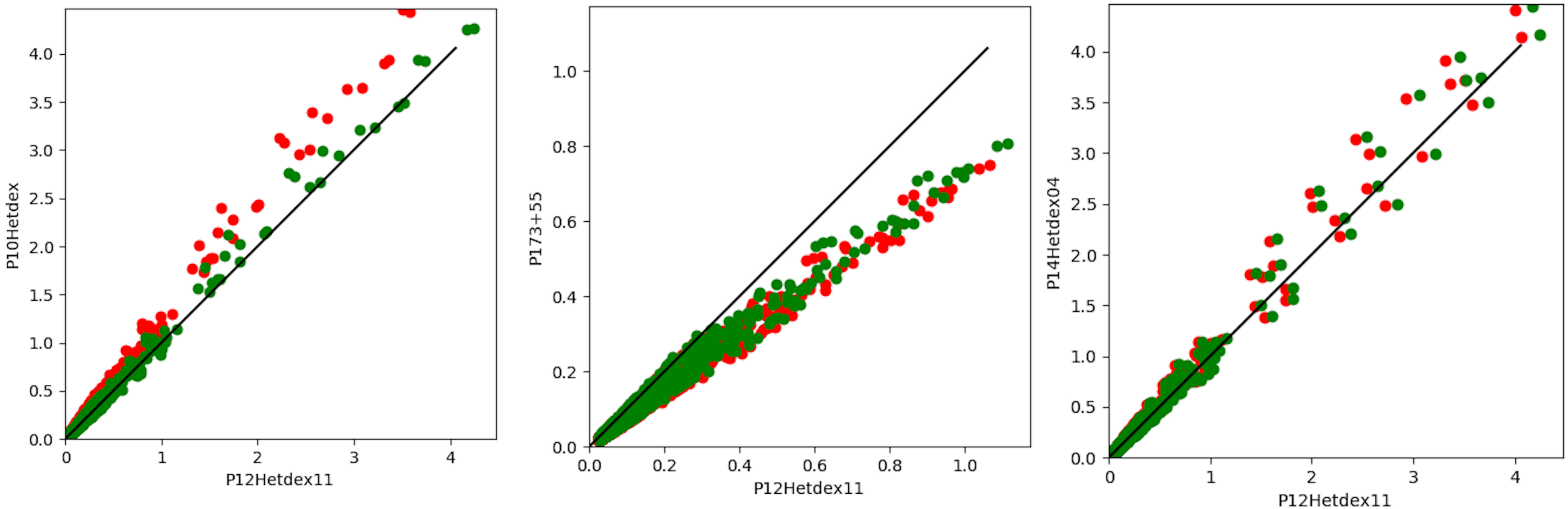
Recovery of completely unmodelled flux has been characterised for various different sizes and fluxes (approximately similar to faint extra galactic sources). Whilst some flux is lost, the recovery is higher than 60%. Red is injected and black is recovered.

Improved flux scale

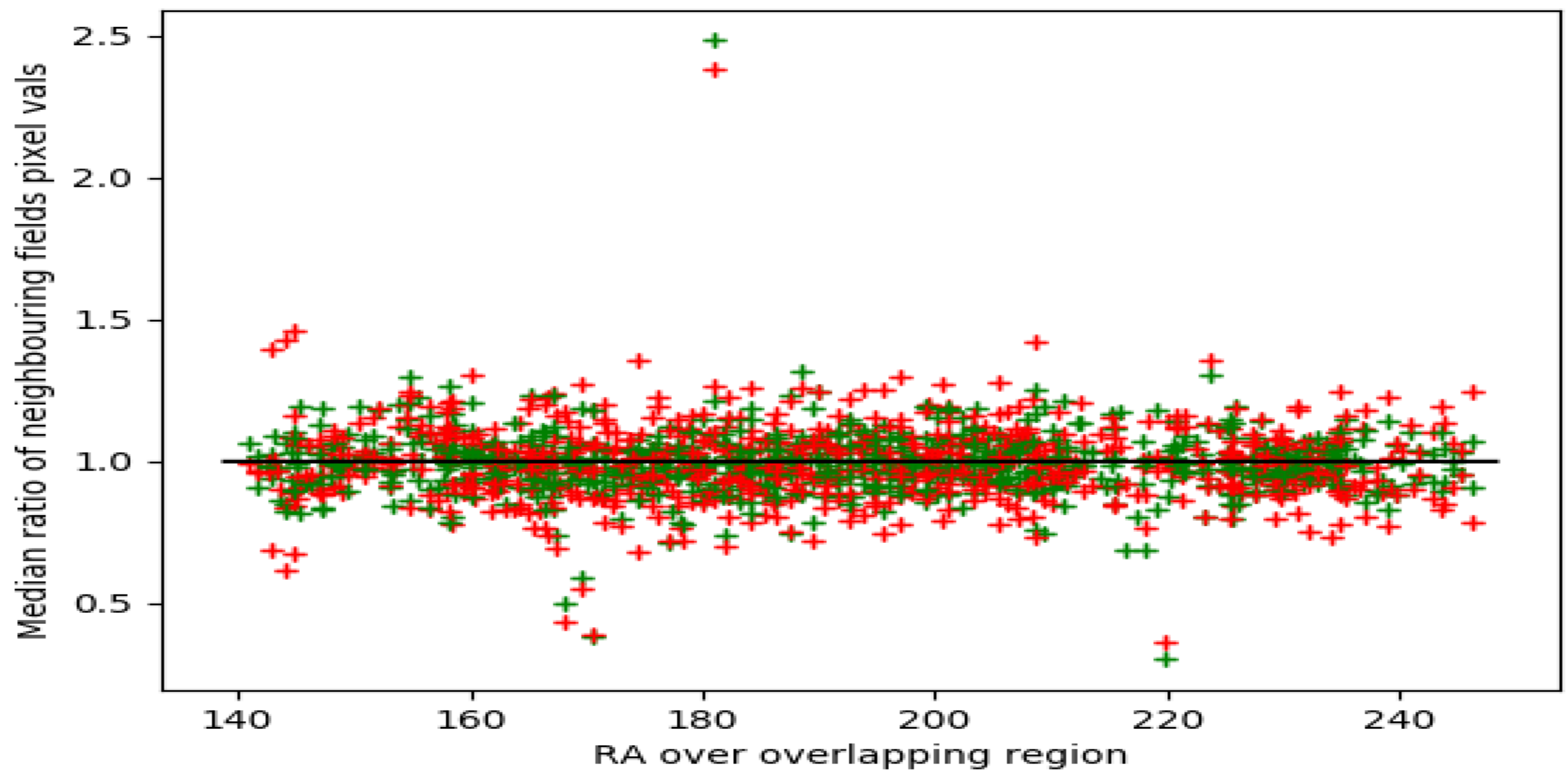


Large overlap between pointings. The values of the pixels with significant signals in these overlapping regions can be used to align the flux scales of the images. Once aligned the entire region can be scaled to match e.g. TGSS, 7C or anything else.

Improved flux scale



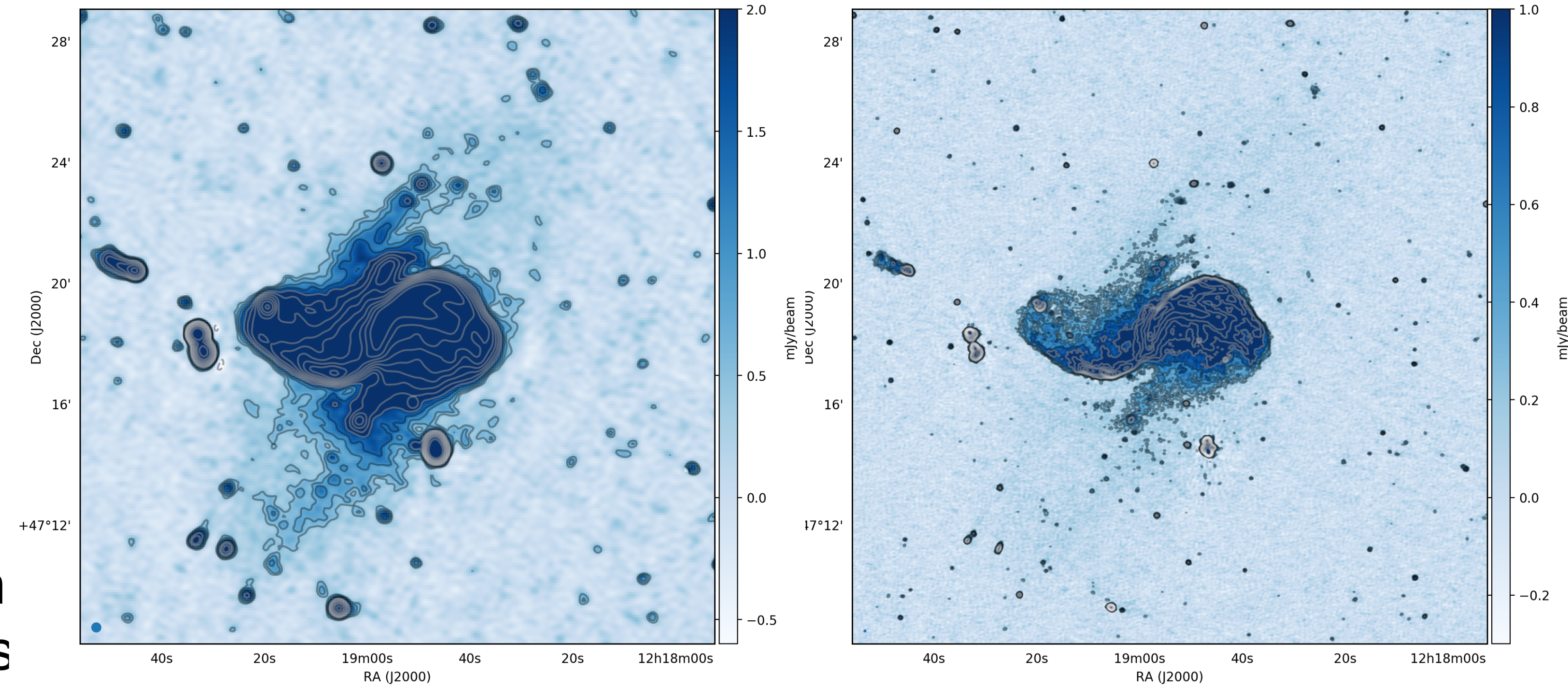
Using an MCMC approach to align (red before alignment and green afterwards) and we are able to decrease the scatter of fluxes a bit but even after this has converged there is still a fair bit of scatter.



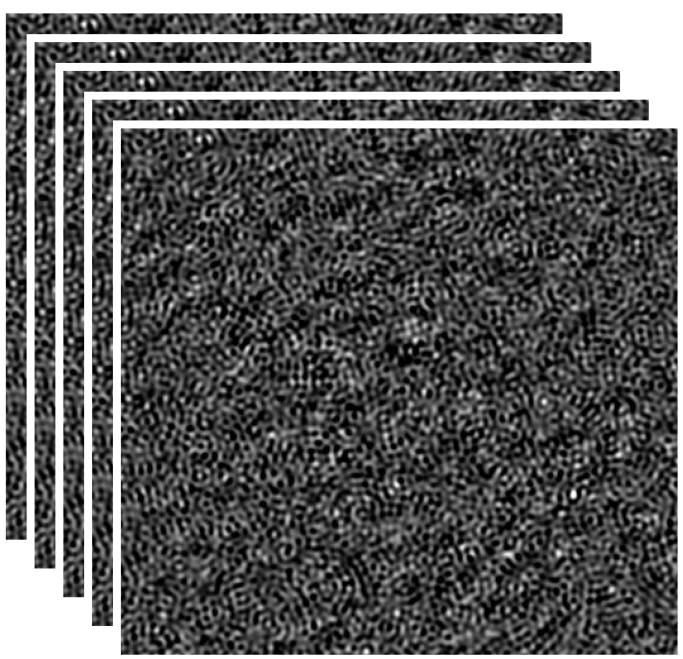
Improvements in LoTSS-DR2

Many data products:

- 6" resolution Stokes I image
- 20" resolution Stokes I image
- 3 channel images over band
- 20" resolution Stokes V image
- 20" resolution Stokes QU cubes (480 planes)
- Very low resolution Stokes QU cubes (480 planes)
- Dynamic spectra of targeted sources
- Data calibrated in a particular direction with all other source subtracted (allows easy reimaging, source subtraction etc)
- Catering for SKSP, MKSP and transients KSP.

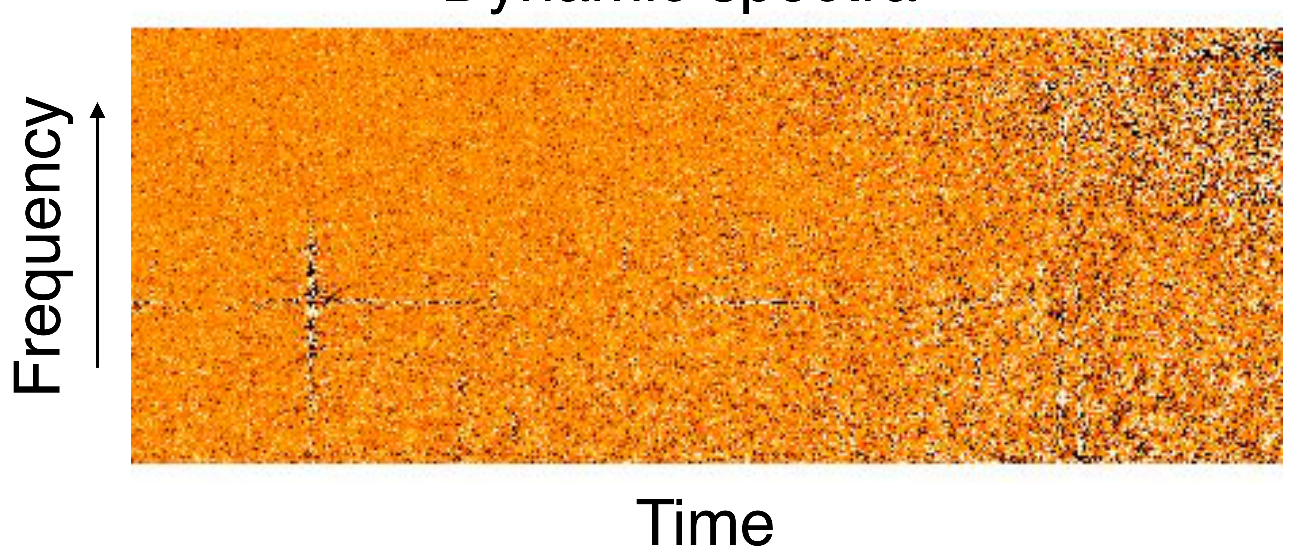


Stokes I images at high (6") and low (20") resolution.
Calibrated data towards targets only ~5GB.



Polarisation cubes

Dynamic spectra

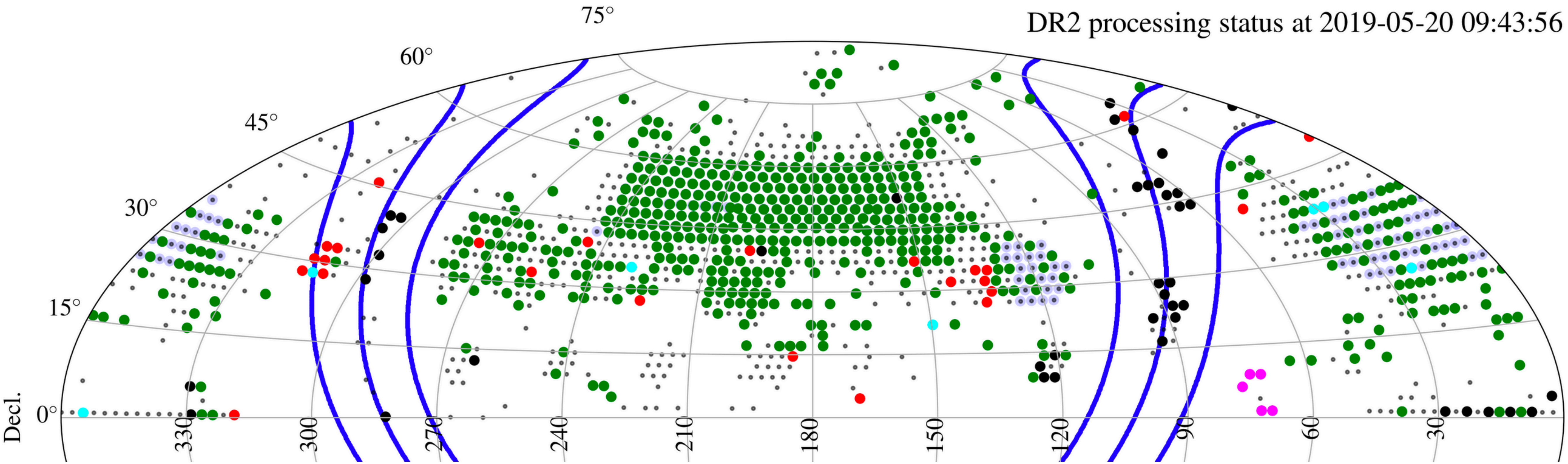


To enable further processing we also keep:

- Facet layout
- Calibration solutions
- Data

DDF-pipeline: Tasse, Hardcastle, Shimwell+

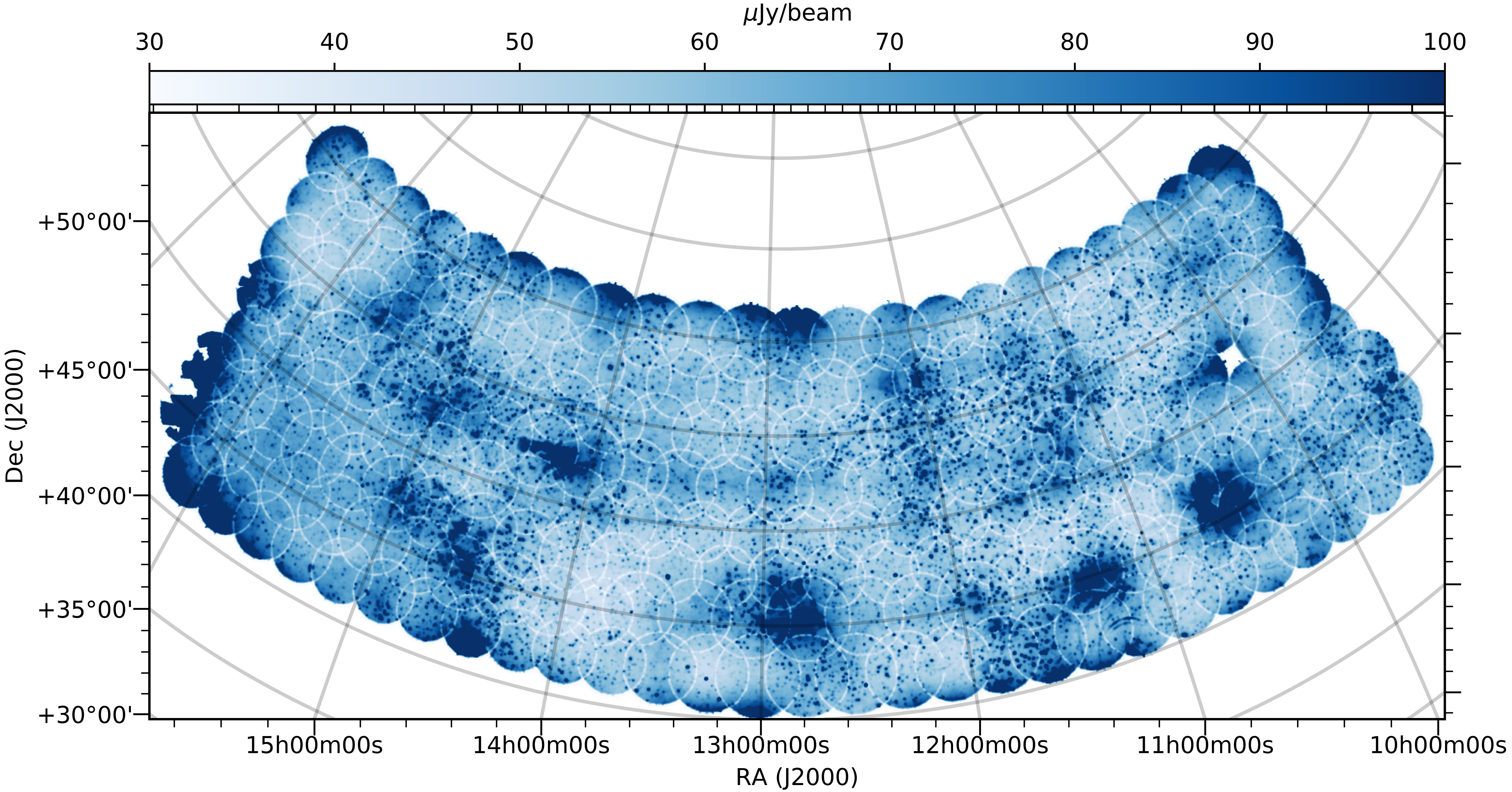
LoTSS data release 2 (LoTSS-DR2)



LoTSS-DR2 will cover around 2,000 square degrees. All is observed and ~2/3 of this area is processed (other areas are processed too) and ~1,800,000 sources are already detected in this region. Outside this region a further ~1,000,000 sources are detected.

(green fully processed, blue direction independent calibration complete, red observed but not calibrated)

LoTSS-DR2



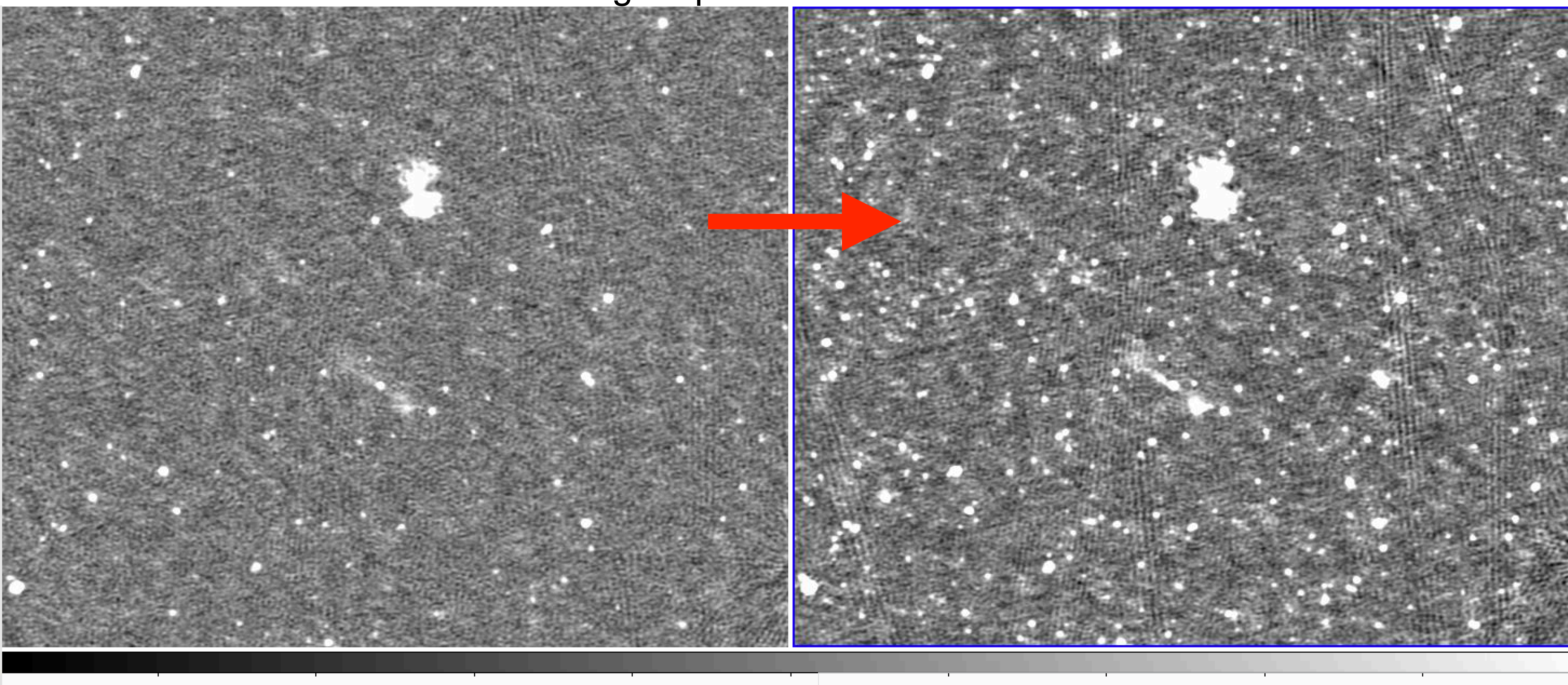
The largest LoTSS-DR2 region covers about 1,200 square degrees and contains 1.4million sources. The other two smaller regions will contain ~800,000 more sources.

The largest region will be complete at the end of May and the mosaics and catalogues will be complete and distributed to the entire collaboration soon after.

LoTSS deep fields

Over ~150hrs gathered towards several important extragalactic fields (Lockman hole, Elais-N1, Bootes so far) and processing of 100s of hrs of data in the region of the north celestial pole in collaboration with the LOFAR EoR group. See Best talk for details on deep fields.

Tests in Elais-N1 indicate we can get optical identifications for ~98% of the radio sources.

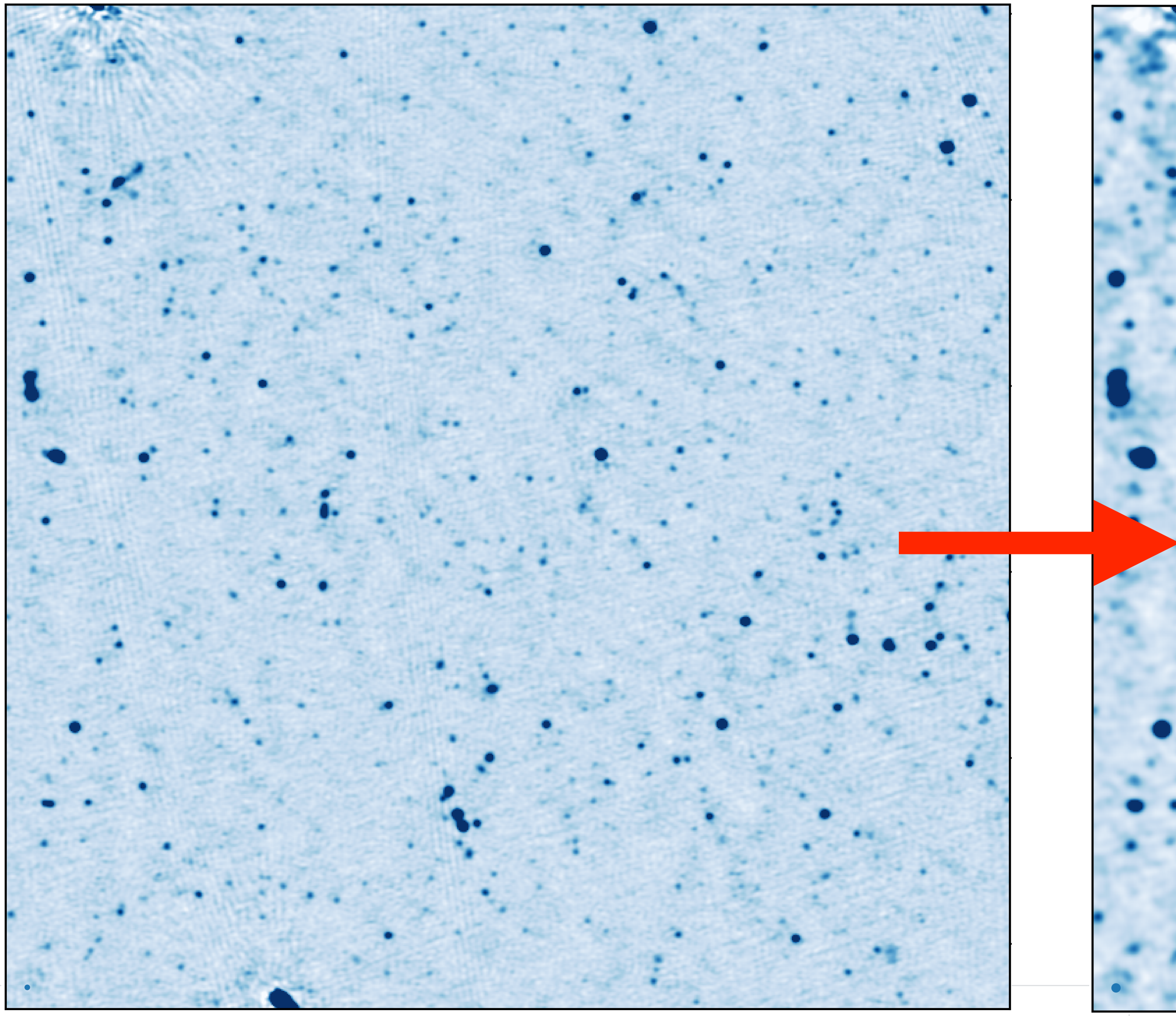


LoTSS depth (~100microJy/beam) — 1000 sources per square degree.

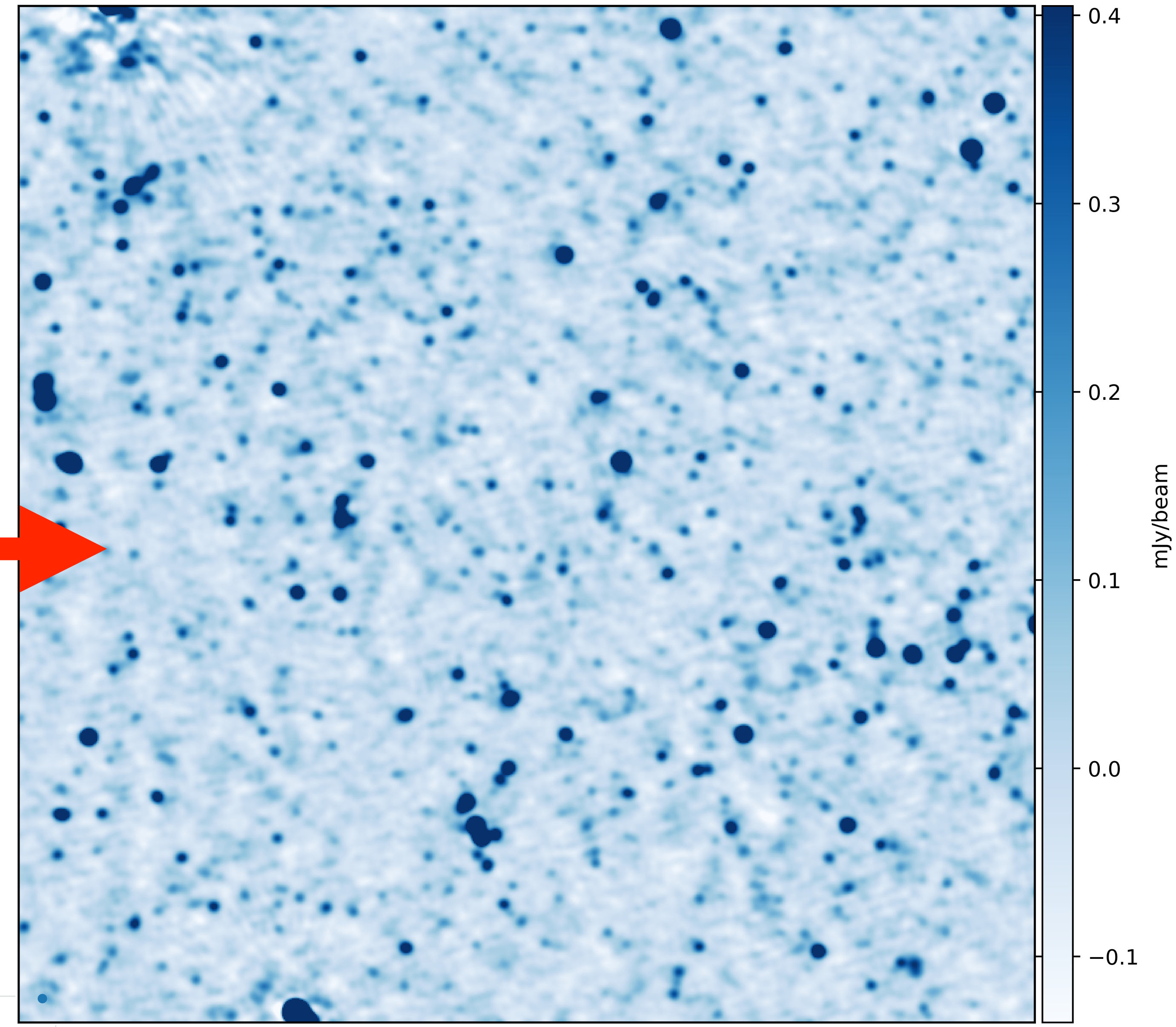
Deep field (~27microJy/beam) — 4200 sources per square degree in the most sensitive region. >50,000 sources in field.

LoTSS deep fields

LOFAR



SKA?



LOFAR deepish field (~27microJy/beam). Just 100hrs observations and 2 fields can be observed simultaneously. LOFAR upgrades mean 4 fields can soon be observed simultaneously.

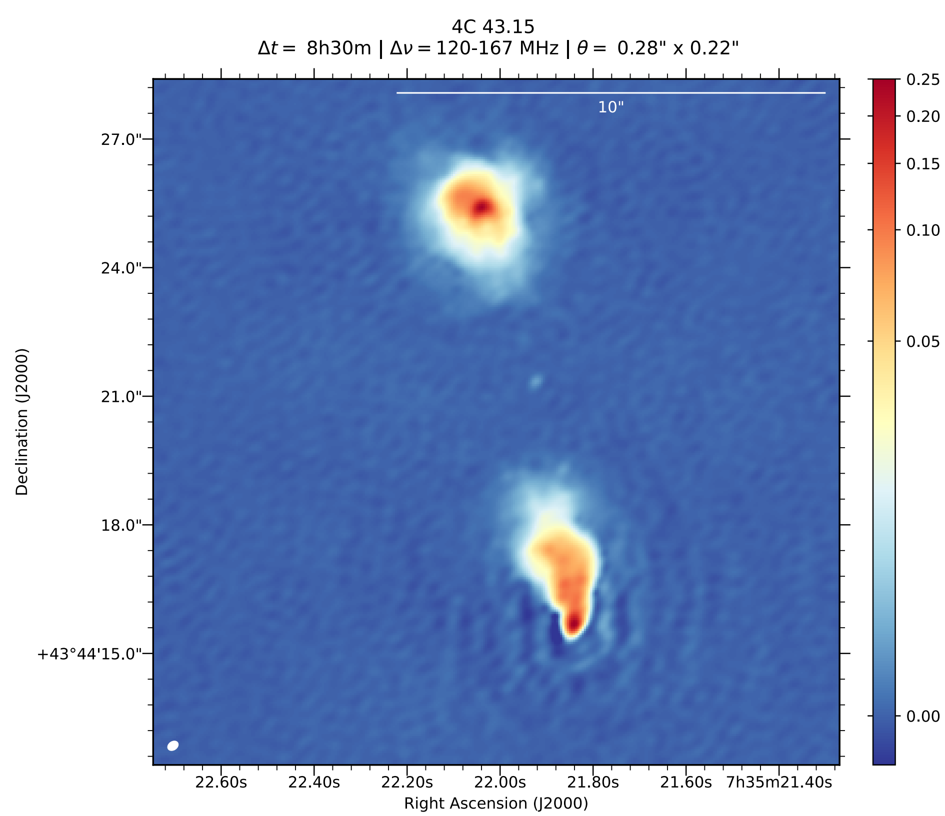
Convolve to 10” to get SKA 150MHz all sky survey image? Approaching confusing limited.

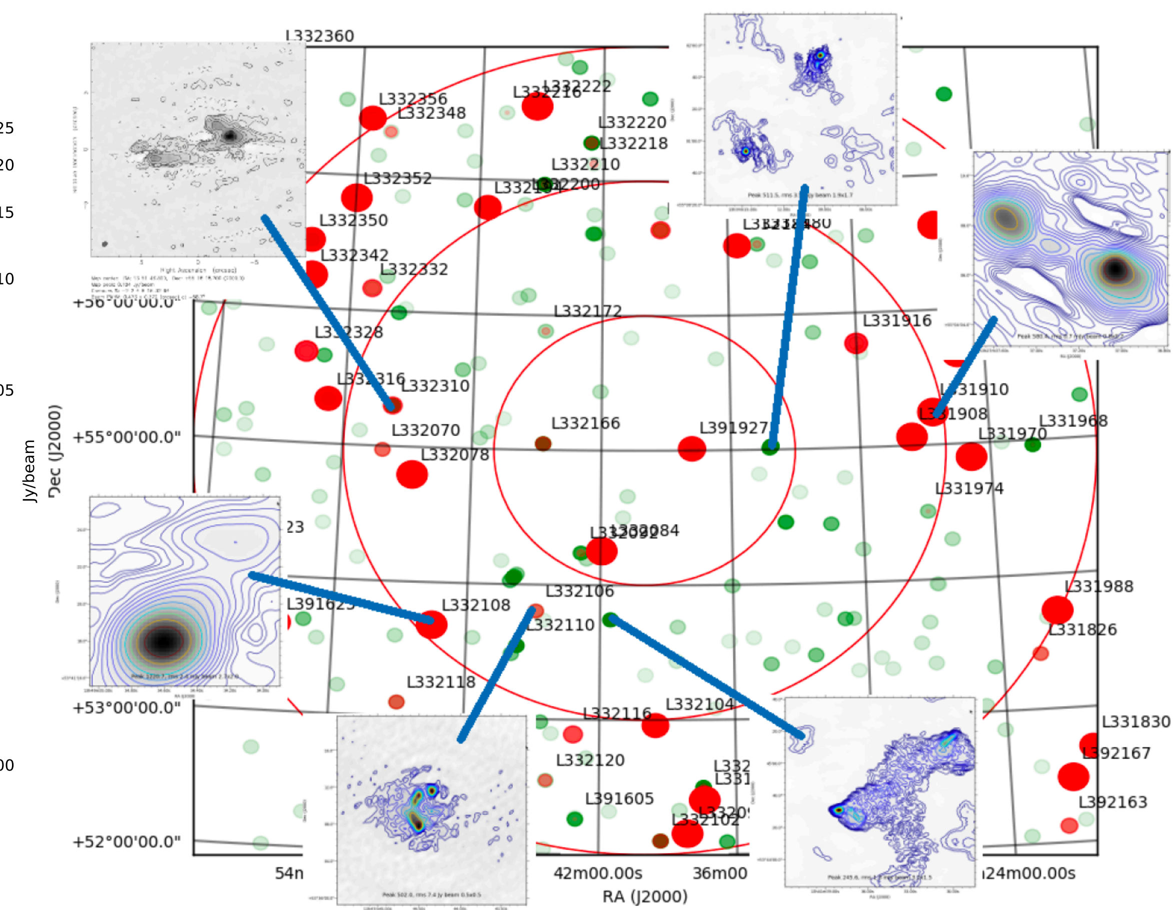
Furthering the LOFAR surveys

Optical followup — WEAVE-LOFAR (Smith+ 2016) will use WEAVE on the WHT and soon begin obtaining spectra for ~a million LOFAR sources.

Radio recombination lines — LoTSS data have sufficient frequency resolution for spectral line work and the data are being analysed to search for RRLs (e.g. Emig+ 2018).

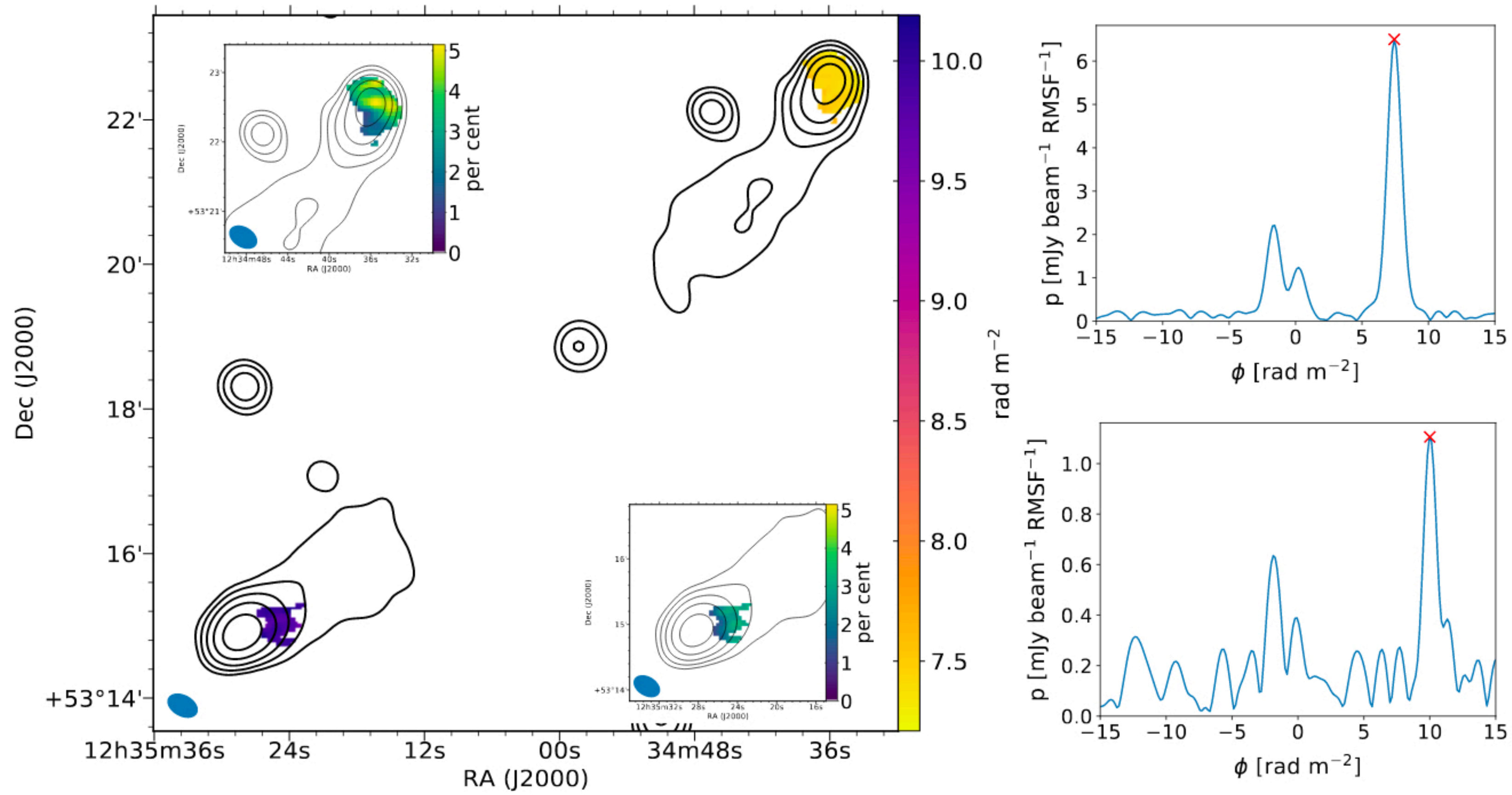
0.3arcsec resolution — LOFAR surveys data are recorded using the full international LOFAR array allowing for 0.3” imaging over the entire surveyed region (images from Sweijen, van Weeren, Jackson, Morabito+)





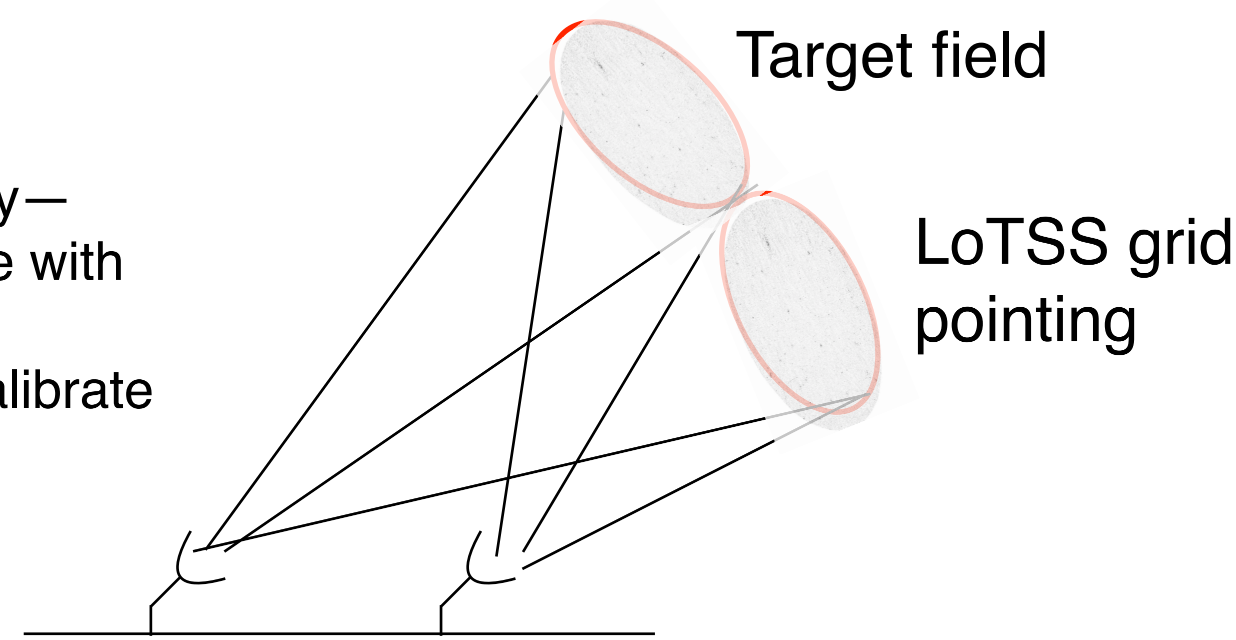
Furthering the LOFAR surveys

Polarisation — LoTSS produces QU cubes and V images. Approximately 1 source per square degree in QU cubes (images from O'Sullivan+ 2018 and also see van Eck+ 2018).

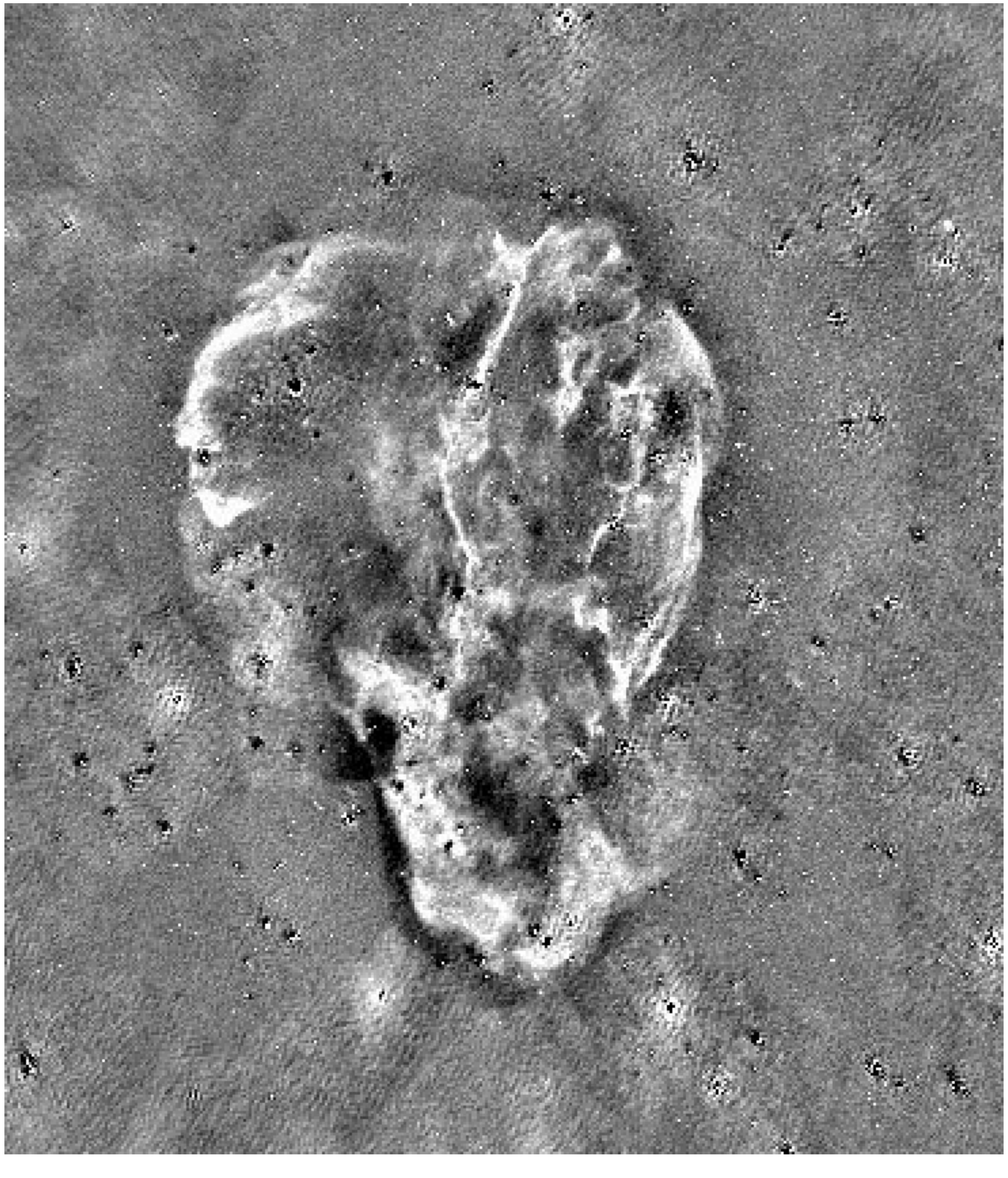


Towards LOFAR 2.0 — better LBA calibration, improved observing efficiency, better long baseline uv-coverage

Working with the observatory— LOFAR users can coobserve with the LOFAR surveys and the surveys team helps users calibrate and image their data.



Furthering the LOFAR surveys



8 degree

Both galactic and low declination fields are difficult — e.g LoTSS images of the Cygnus loop show emission over huge areas that are very hard to accurately model and calibrate against. At low declination the primary beam gets larger, the sensitivity decreases and getting enough signal for calibration becomes harder.

Summary

- The products from the LOFAR HBA survey data release 1 (LoTSS-DR1) covers 400 square degrees and contain 325,000 radio sources and over 200,000 with optical ids and redshifts. The images, software and catalogues completely public since February.
- LoTSS-DR2 will cover ~2,000 square degrees and contain ~2,000,000 sources — making LoTSS-DR2 (10% of the entire LoTSS) the largest radio survey ever conducted.
- LoTSS-DR2 products are made available to the collaboration as soon as they are processed. The first large are LoTSS-DR2 catalogue will be released internally soon (after quality checks).
- There is a lot of exciting science that can be done with the survey data and it has a huge legacy value.