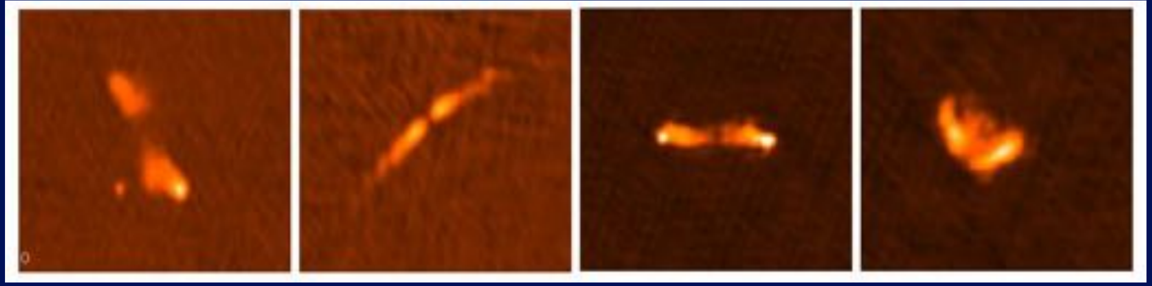
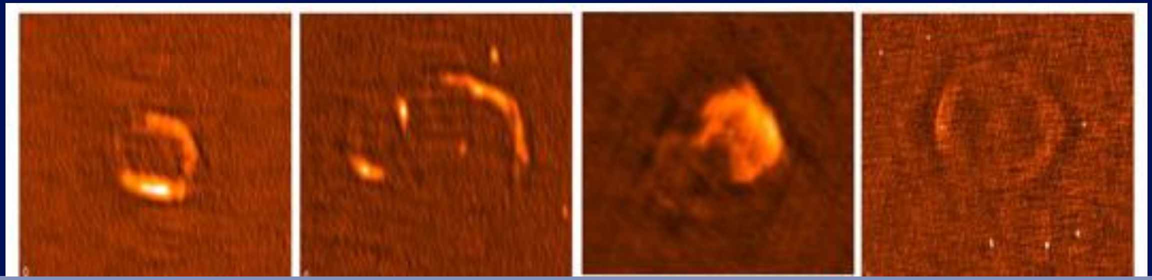


How to SPAM the 150 MHz sky



Huib Intema | Leiden Observatory

04/11/2016

Main collaborators:
Preshanth Jagannathan (UCT/NRAO)
Kunal Mooley (Oxford)
Dale Frail (NRAO)

Talk outline

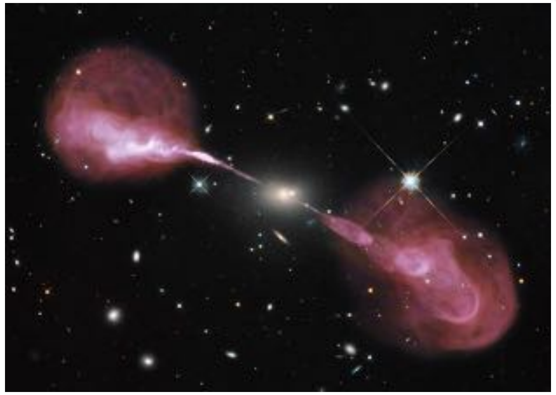
- The need for a low-frequency radio reference survey
- The TGSS survey
- The SPAM pipeline
- Application to TGSS
- TGSS ADR as a reference survey
- Current status and future plans

LOFAR: the Low Frequency ARray

SKA-low: Square Kilometer Array at low frequencies

Scope

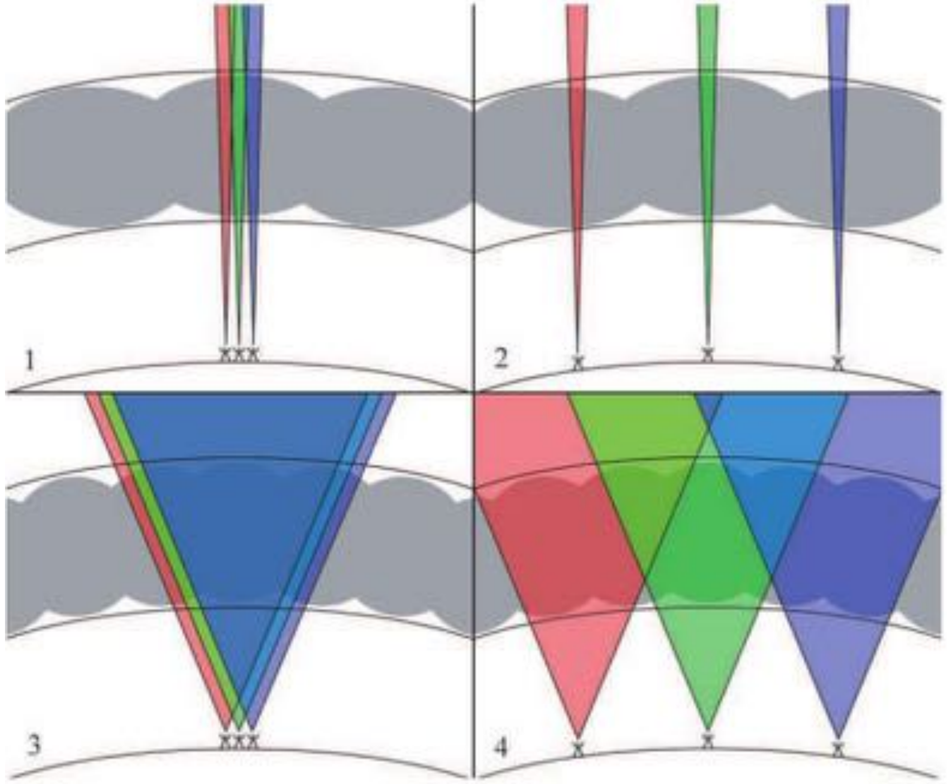
- The Universe is very transparent for long radio waves
- Rich tradition on surveying the sky at low radio frequencies
  - Cambridge catalogs (UK), NRAO surveys (USA), Westerbork (NL), Molonglo (AUS)
- The intrinsic large field-of-view provides a high survey speed
  - But the resolution is typically poor
- Renewed astronomical interest to survey the sky at sub-GHz frequencies
  - Higher resolution, better sensitivity, new technologies (LOFAR, SKA-low)
- Some main science drivers are
  - High-redshift neutral hydrogen (Epoch-of-Reionization)
  - Pulsars and transients (GRBs, FRBs, GRWs, …)
  - Exo-planets
  - Galaxy cluster formation and evolution
  - Cosmic magnetism



Bonafede+ 2014

Scope

- Direction-dependent effects are a major problem for wide-field, low-frequency radio surveys
  - Ionospheric dispersive delay and Faraday rotation
  - Complex antenna/station beam patterns
- Main driver behind recent development of direction-dependent (DD) calibration schemes
  - field-based calibration, SPAM, MeqTrees, Sagecal, LOFAR facet calibration, KillMS, …
- Having a good reference sky model at similar frequency and resolution is crucial
  - Local astrometry is not conserved due to DD ionospheric phase gradients
  - Complex antenna beam patterns introduce uncertainty in measured flux densities
  - Large uncertainty in low-frequency flux density scale in general

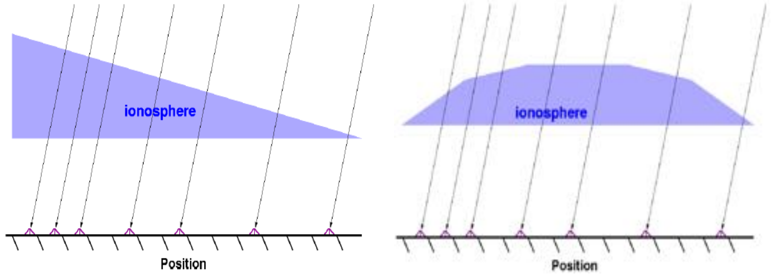


Intema+ 2009

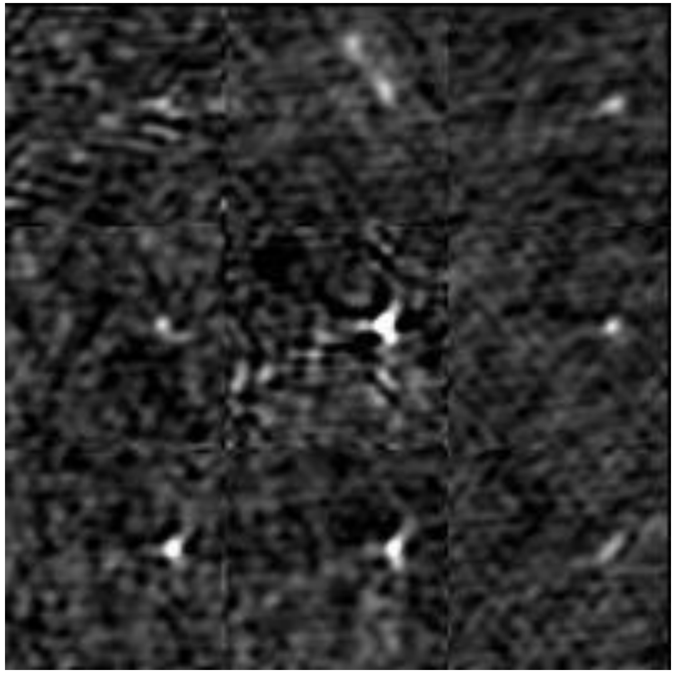
Scope

Major distortion is ionospheric dispersive delay (similar to optical seeing)

$$\phi = \frac{e^2}{4\pi\varepsilon_0 m}\lambda\int_0^d n_e(s)\,ds$$



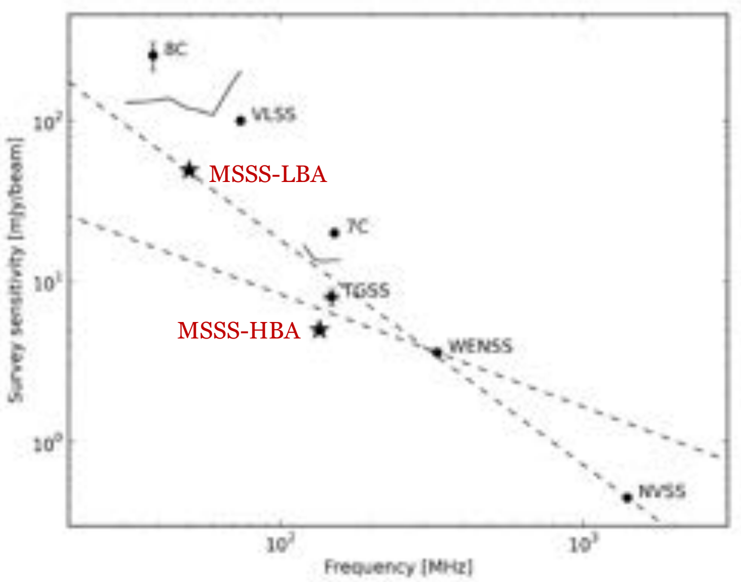
- A radio interferometer measures phase differences, therefore senses the differential structure in the electron column density (TEC)
  - TEC gradients cause apparent source shifts
  - Higher TEC structures cause source distortions
  - TEC structure varies with time and direction
  - Very sensitive: 1 radian per 0.01 TECU at 75 MHz
- Time series of 1-minute snapshot images of 9 sources distributed over a single 10-degree field-of-view of the VLA at 74 MHz
  (movie created by W.D. Cotton, NRAO)

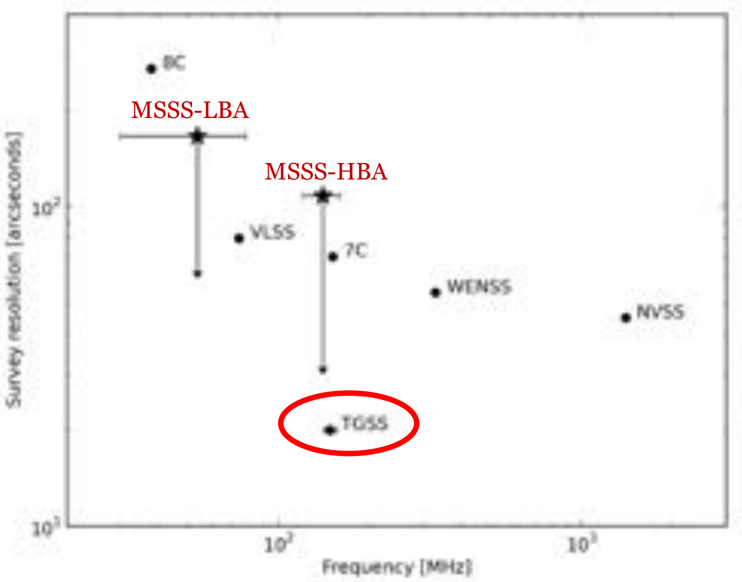


Scope

Selection of available/upcoming reference sky surveys

Survey	Frequency	Sensitivity	Resolution	Area
MSSS-LBA	30–78 MHz	$\lesssim$50 mJy beam^{-1}	$\lesssim$150″	20 000 □° ($\delta > 0°$)
8C	38 MHz	200–300 mJy beam^{-1}	4.5′ × 4.5′ csc(δ)	3000 □° ($\delta > +60°$)
VLSS	74 MHz	100 mJy beam^{-1}	80″	30 000 □° ($\delta > -30°$)
MSSS-HBA	120–170 MHz	$\lesssim$10–15 mJy beam^{-1}	$\lesssim$120″	20 000 □° ($\delta > 0°$)
7C	151 MHz	20 mJy beam^{-1}	70″ × 70″ csc(δ)	5500 □° (irregular coverage)
TGSS	140–156 MHz	7–9 mJy beam^{-1}	20″	32 000 □° ($\delta > -30°$)
WENSS	330 MHz	3.6 mJy beam^{-1}	54″ × 54″ csc(δ)	10 000 □° ($\delta > +30°$)
NVSS	1400 MHz	0.45 mJy beam^{-1}	45″	35 000 □° ($\delta > -40°$)



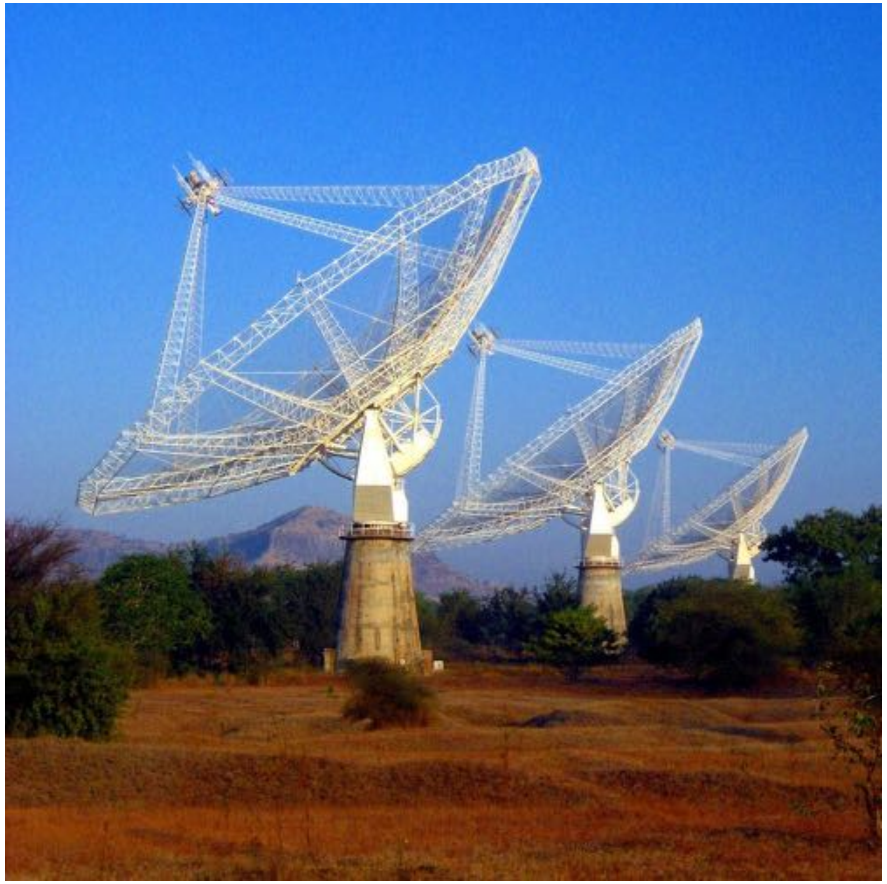


Heald+ 2015

Overview of the TGSS

TIFR GMRT Sky Survey

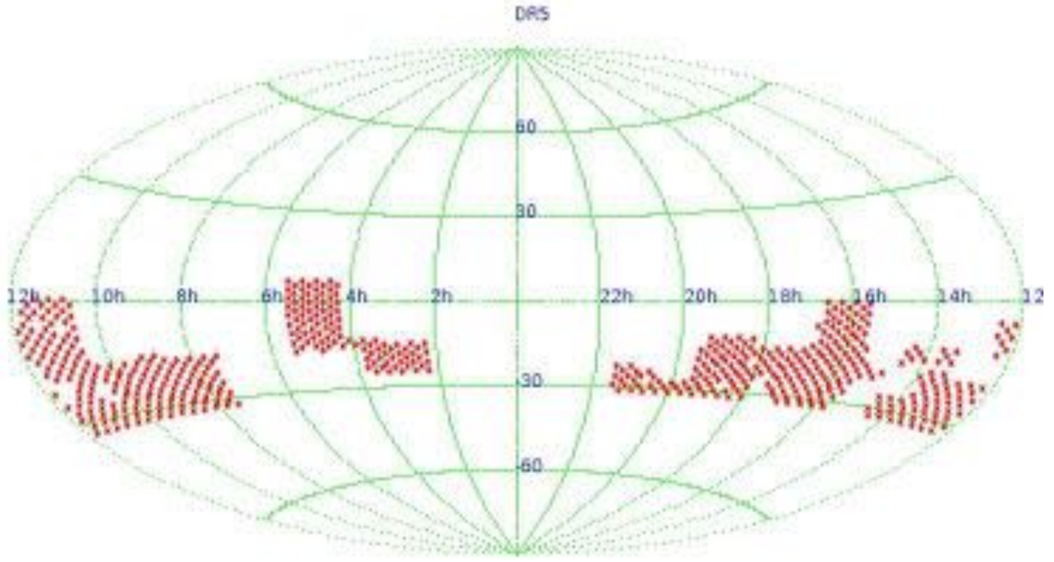
- Continuum survey at 150 MHz with the GMRT
- PI-driven project
- 16 MHz bandwidth, 20” resolution, ~3 degree FoV
- 5,336 pointings covering DEC +90 to -55 degrees
  37,000 square degrees = 90 percent of the radio sky
- Observing grid following FIRST scheme
- 15 minutes/pointing, median 5-7 mJy/beam RMS
- 2,000 hours granted and observed between 2010-2012
- http://tgss.ncra.tifr.res.in/



Overview of the TGSS

TGSS data processing and releases

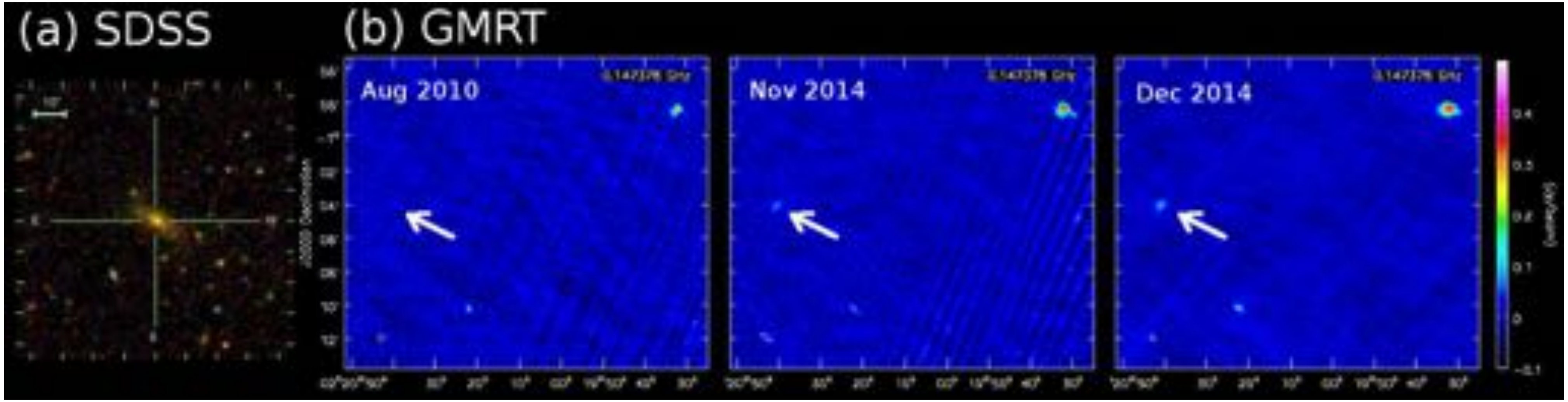
- Data processing based on old AIPS++ pipeline, selfcal only
- Dedicated 100-node compute cluster at NCRA
  - Recently upgraded to 1600 cores, 80 TB RAM, 1 PB diskspace
- 5 data releases to date, last one (DR5) in late 2012
- DR5 contains about 10 percent of the survey area
- New releases promised, but current state (still) unclear
- Pilot and main survey remain unpublished
- Website not updated since 4 years
- All raw data has become publicly available through GMRT archive (https://naps.ncra.tifr.res.in/goa/)



SPAM pipeline

Pipeline development triggered by GEMS survey

- GMRT Exploration of the transient Meterwavelength Sky
- Slow transient survey in STRIPE82 region (PI Mooley, Oxford)
- 150 MHz, 300 deg², 2 epochs
- 4-7 mJy/beam rms noise, 20” spatial resolution
- Uses TGSS data as extra epoch



SPAM pipeline

Source Peeling & Atmospheric Modeling

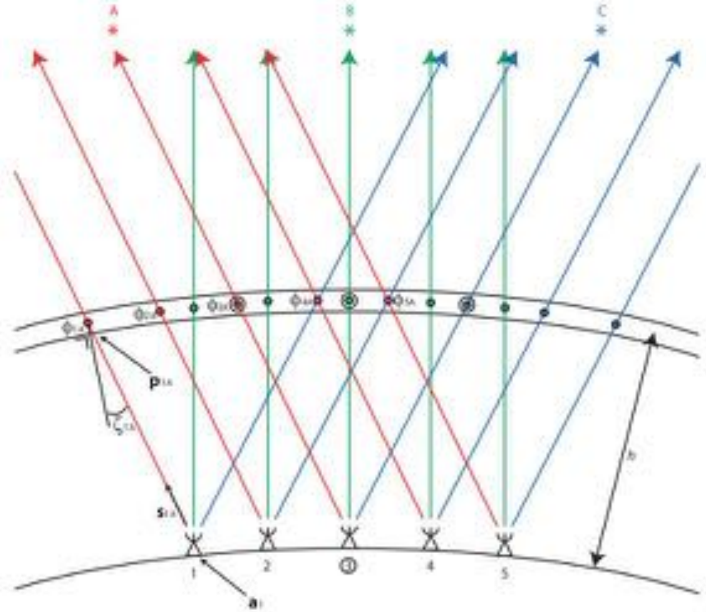
- Python/AIPS-based software for in-beam ionospheric calibration
  Relies heavily on ParselTongue
- Performs wide-field (direction-dependent) ionospheric calibration, modeling, and imaging
- Strategy and algorithms developed and improved for VLA and GMRT (since 2007)
- Developed into fully automated pipeline for GEMS project (late 2014)
- Custom-build scheduler for parallel batch job processing on Linux compute cluster (thanks NRAO AOC computing staff)

- Pipeline processing worked extremely well for both GEMS and TGSS data
- "Small" steps from STRIPE82 to DR5 (demonstrator) to whole TGSS survey

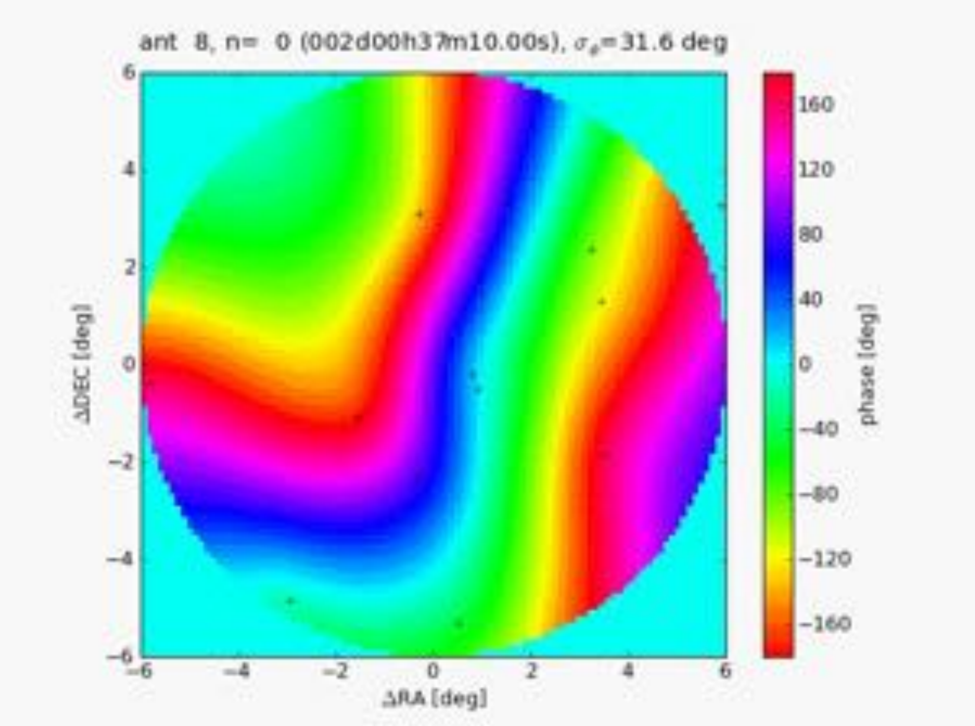
SPAM pipeline

SPAM core functionality

- A measurement of the local ionospheric TEC structure is obtained by phase calibrating on bright sources within the field-of-view (e.g., peeling)
- The measured phases of all source-antenna pairs can be mapped onto ionospheric layer
- All phases per time interval are fitted with a single model (based on thesis work by Van der Tol, 2009)
- Model predicts phases corrections in arbitrary directions for imaging full field-of-view
- Example time series of a dual-layer phase screen model for narrow-band VLA 74 MHz observation
  - Phase screens fitted each 10 sec to peeling phases of ~10 sources



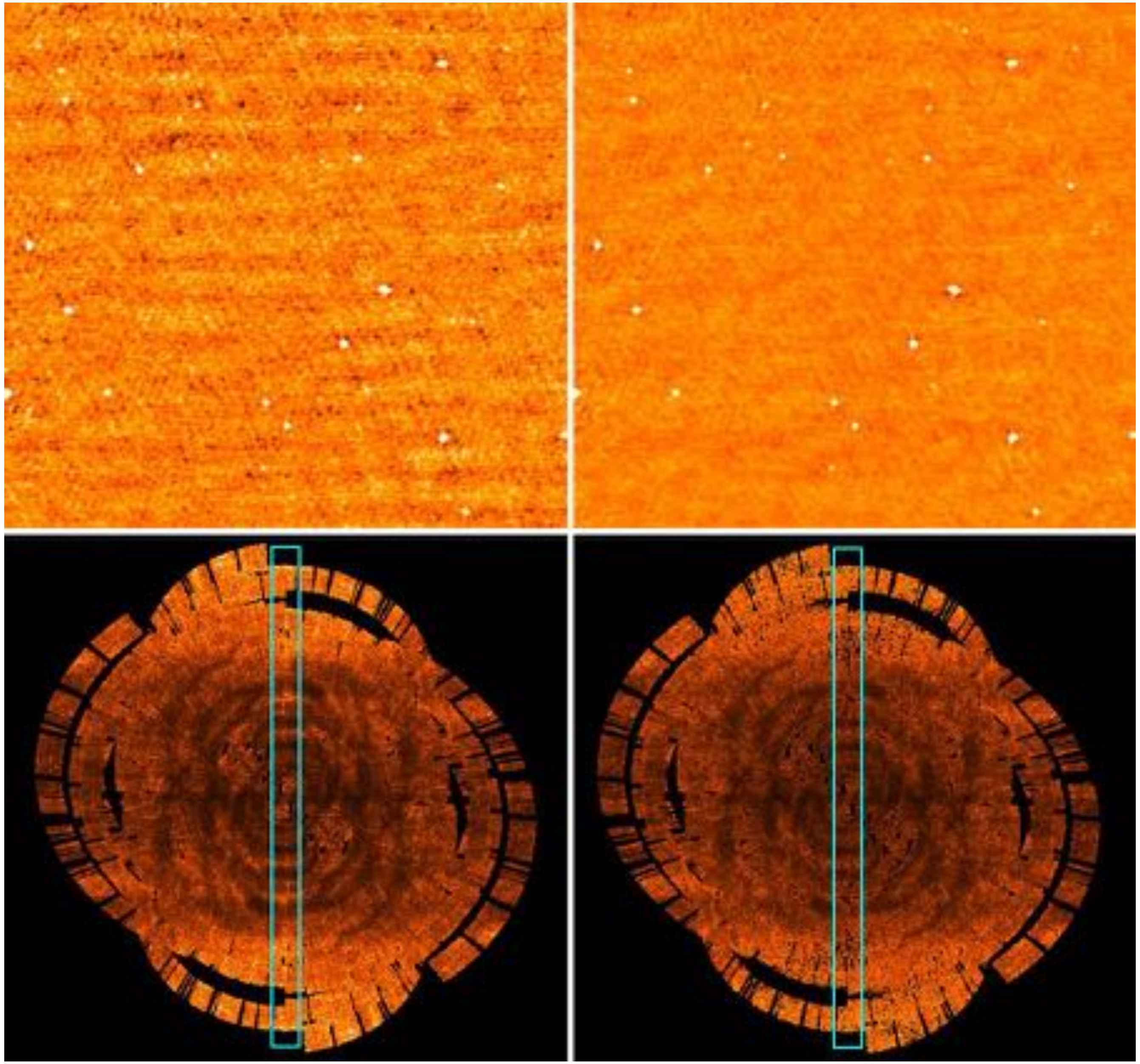
Intema+ 2009



SPAM pipeline

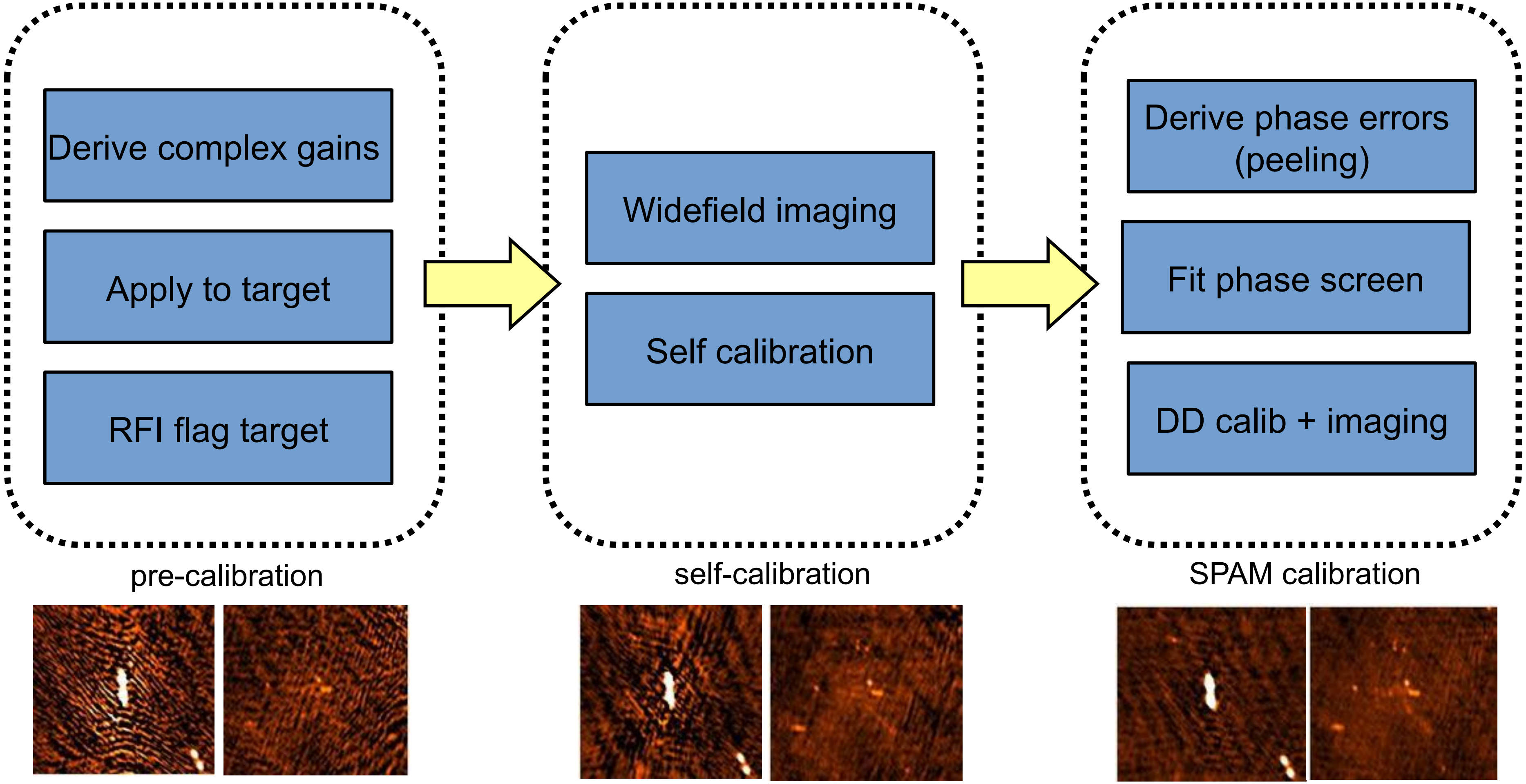
Image-based flagging

("ripple killer")



SPAM pipeline

Functional overview



Application to TGSS

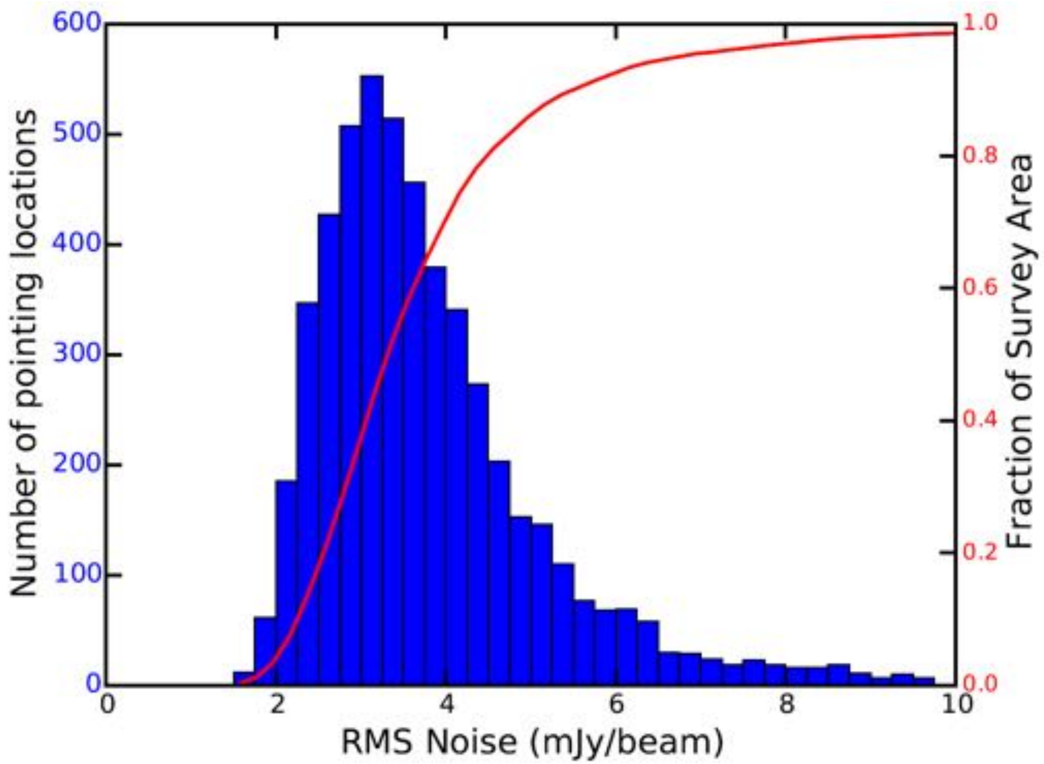
SPAM pipeline performance

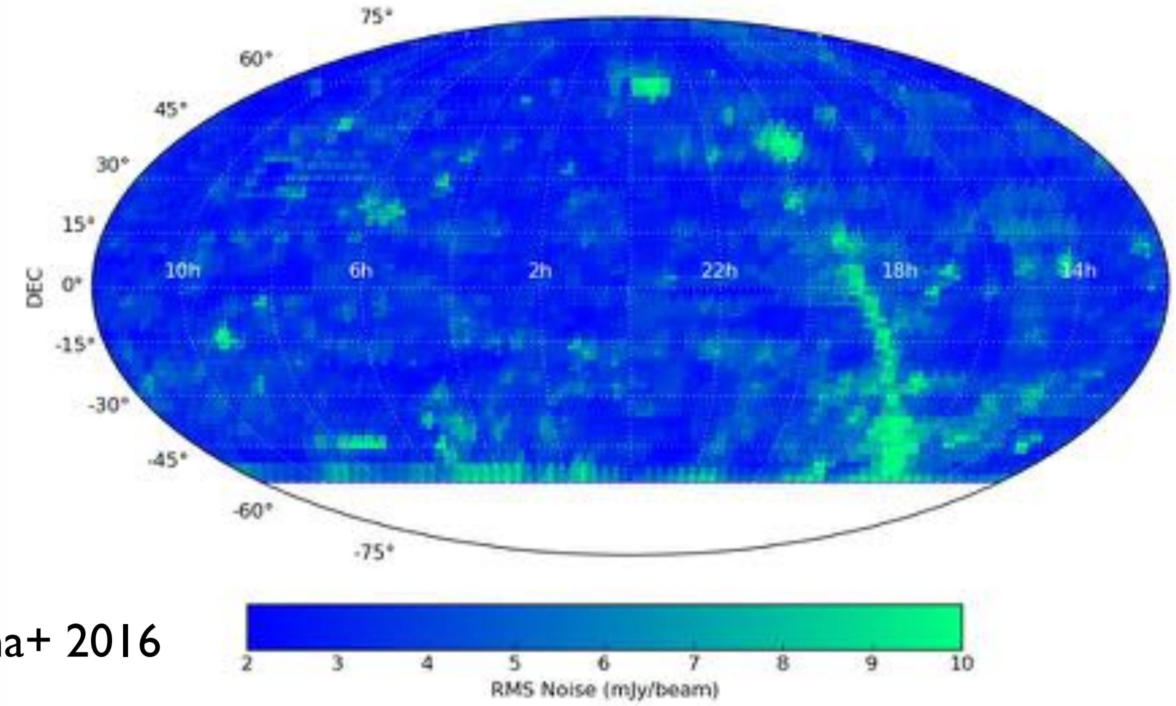
- Total time-averaged raw data volume 1.8TB
- Data conversion & pre-calibration: 30 minutes/observation
  - 200 observations = 100 CPU hours
- SPAM pipeline: 3 hours/pointing
  - 5,500 pointings = 16,500 CPU hours = 1.9 CPU years
  - Comparison LUSTRE vs SSD vs HD vs RAM drive: RAM drive by far best performance
    Requires minimization of temporary data storage
- Parallel SPAM processing: 12 jobs/node, 4 nodes = 2 weeks(!)
- With 2 passes, 98 percent of the TGSS survey area was processed successfully
- Remaining 2 percent were problematic data and sky areas, and required manual work
- Several post-imaging corrections introduced to ensure internal flux consistency

Application to TGSS

Sensitivity distribution

- 80 percent of the pointing images have a noise level between 2-5 mJy/beam
- Higher noise mostly in Galactic plane, near bright sources (Cas A, Cyg A), and at very low DEC

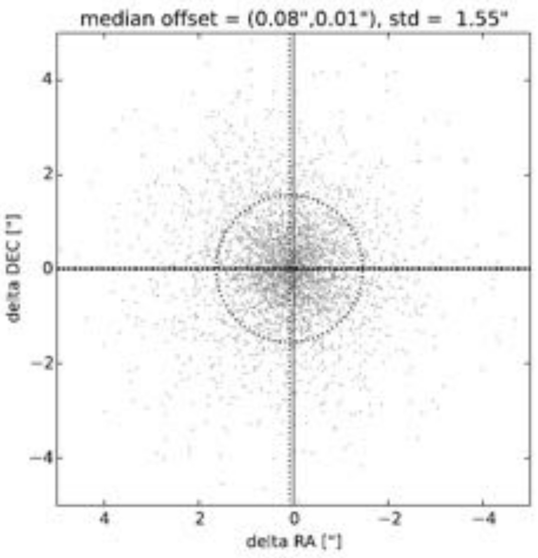


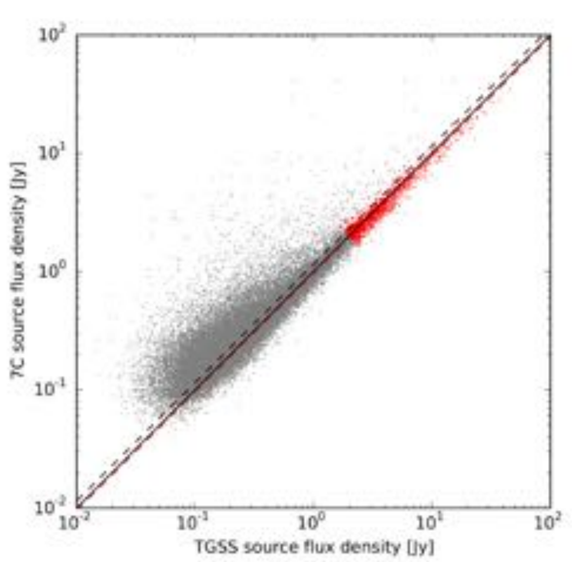


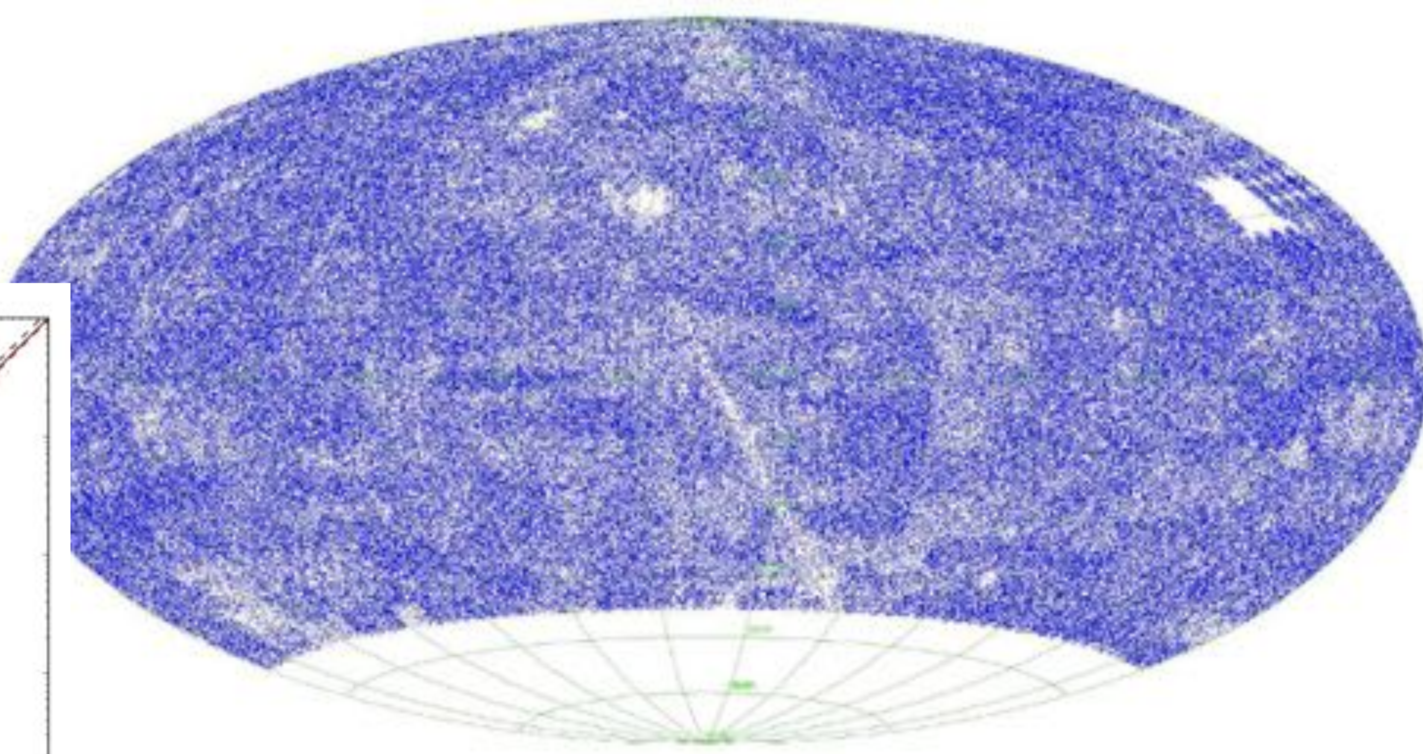
Intema+ 2016

Application to TGSS

- 0.62 Million radio sources detected at 7-sigma level
- Source density correlates with background noise
- Majority of sources are unresolved at 25”
- Positional accuracy < 2”
- Flux density accuracy ~10 percent*







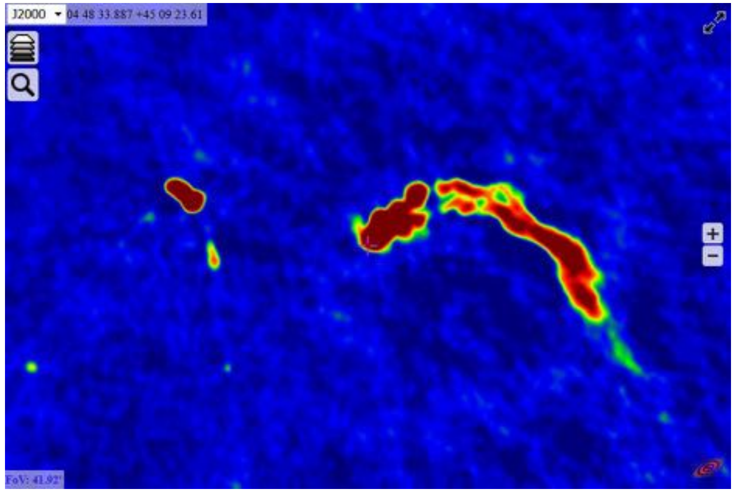
Intema+ 2016, A&A (accepted), arXiv:1603.04368

Application to TGSS

- First full data release of the GMRT 150 MHz sky survey in March (TGSS Alternative Data Release)
- Essential low-frequency reference survey at 25” resolution and 2-5 mJy/beam noise
- Covers 90 percent of radio sky, nearly complete above -53° DEC (significant overlap with LOFAR, MWA and SKA)
- Fully automated processing pipeline including (SPAM) DD ionospheric calibration
- Pilot project for LOFAR surveys products on ASTRON VO (vo.astron.nl)

http://tgssadr.strw.leidenuniv.nl

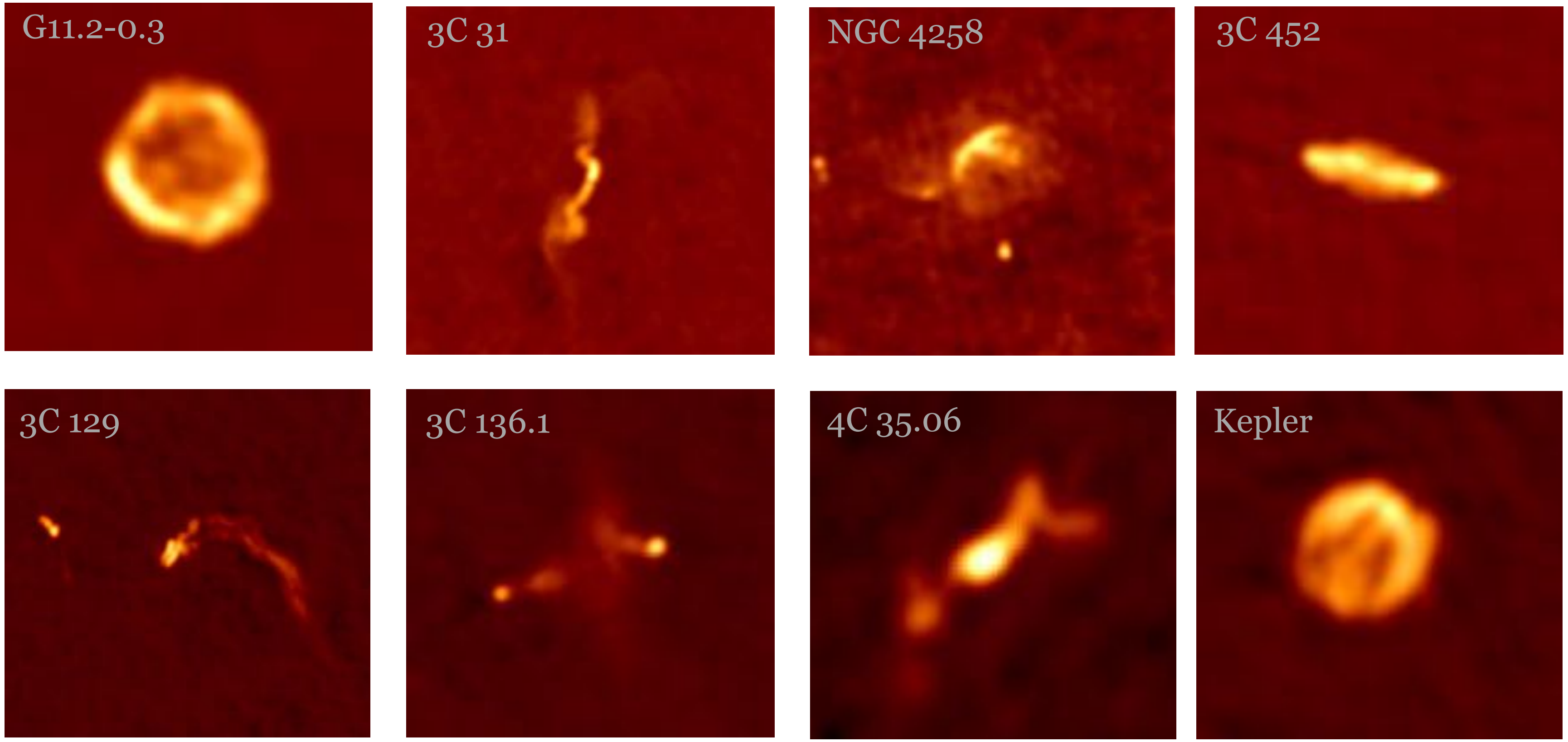
Interactive access through CDS Aladin



Mosaicking options through NASA SkyView

Application to TGSS

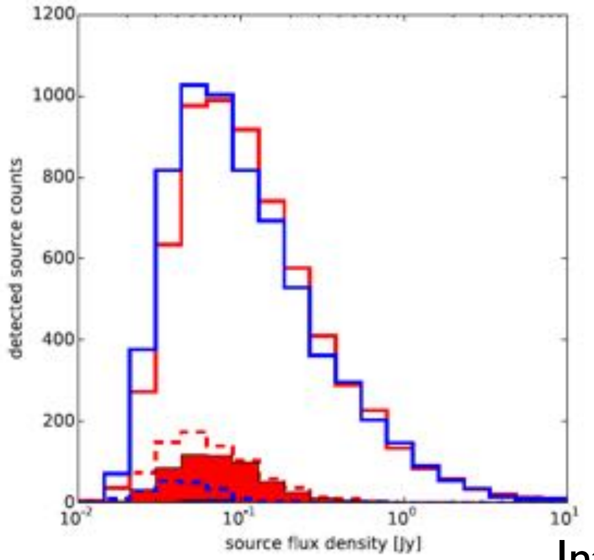
Pretty image gallery



Application to TGSS

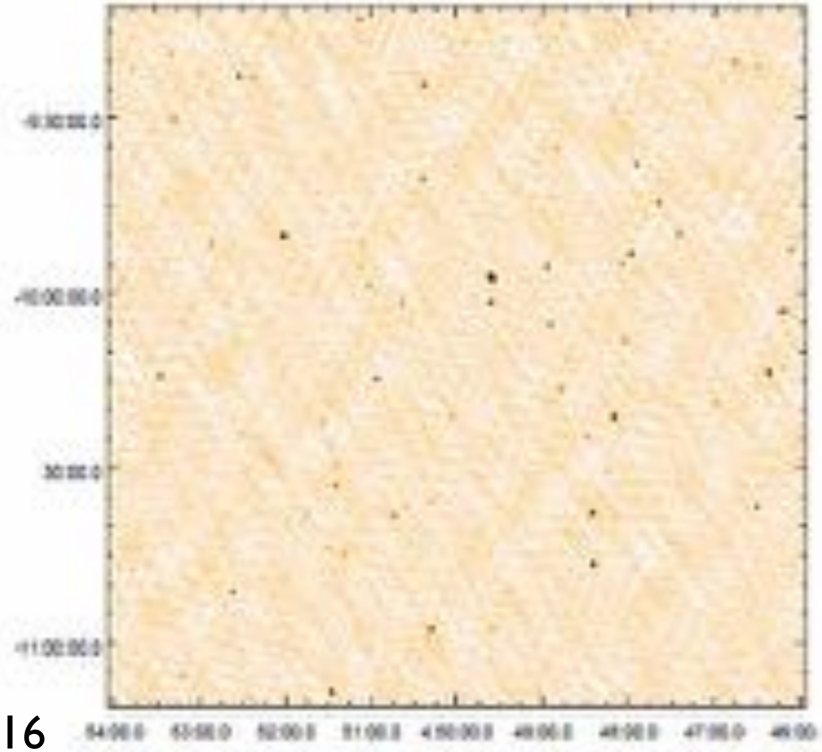
Typical example of SPAM and original TGSS DR5

- Reduction of overall background noise (2-5 mJy/beam versus 5-9 mJy/beam)
- Reduction of artifacts around bright sources
- Increase of peak fluxes
- Improvement of image fidelity (fewer false detections)

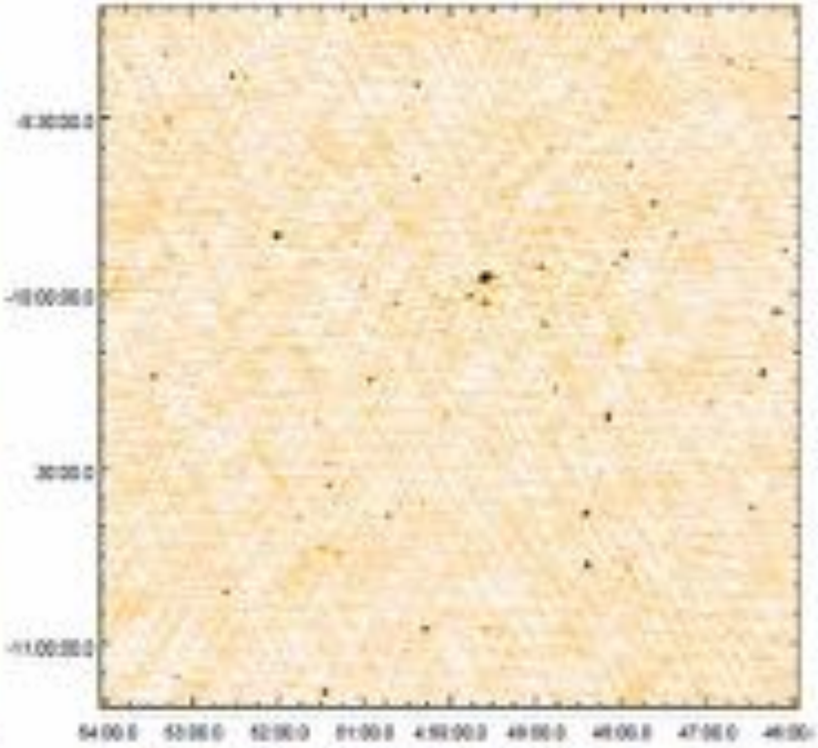


Intema+ 2016

SPAM processing

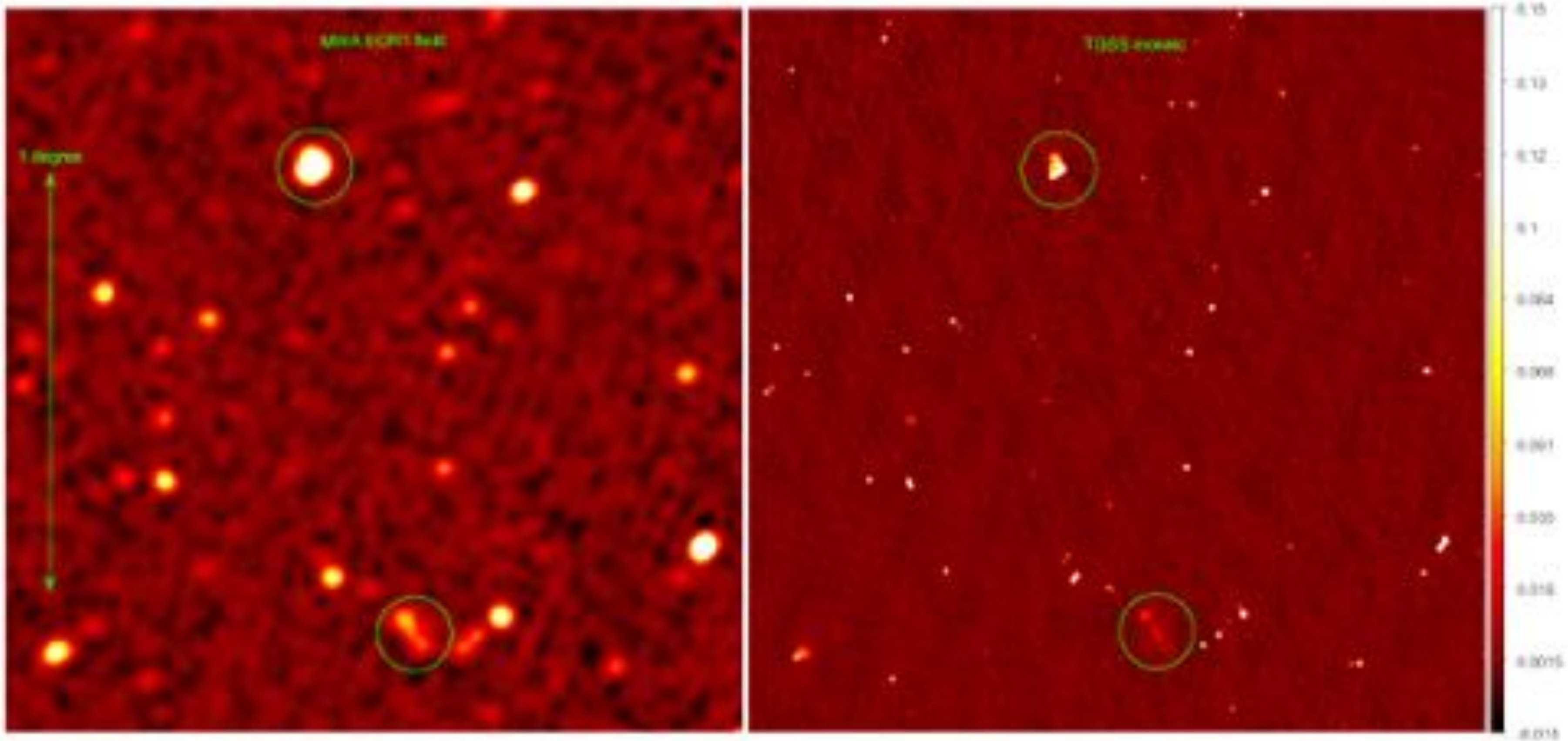


original processing



TGSS as a reference survey

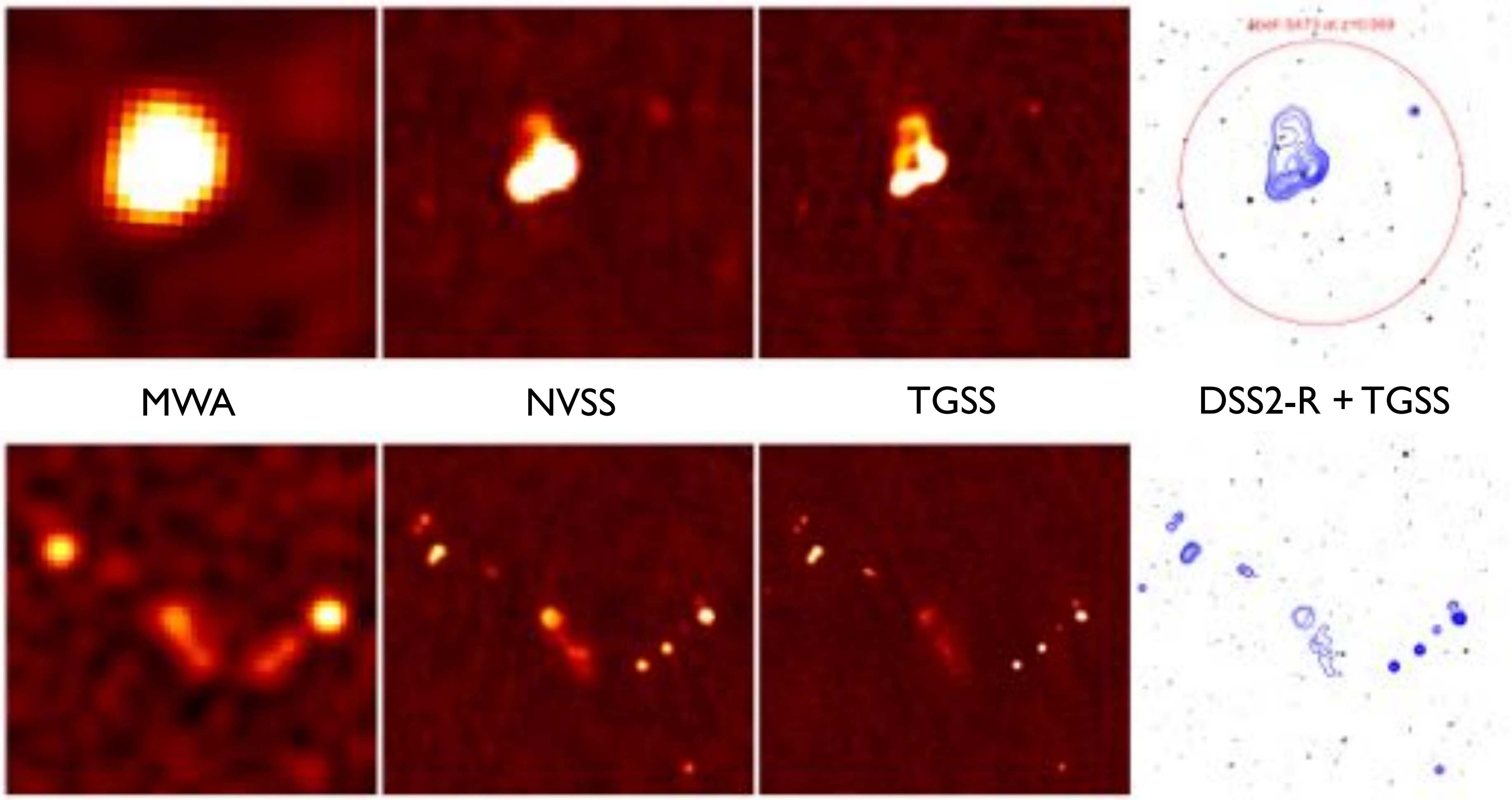
Comparison against MWA (no long baselines)



Hurley-Walker+ 2016
Wayth+ 2015

TGSS as a reference survey

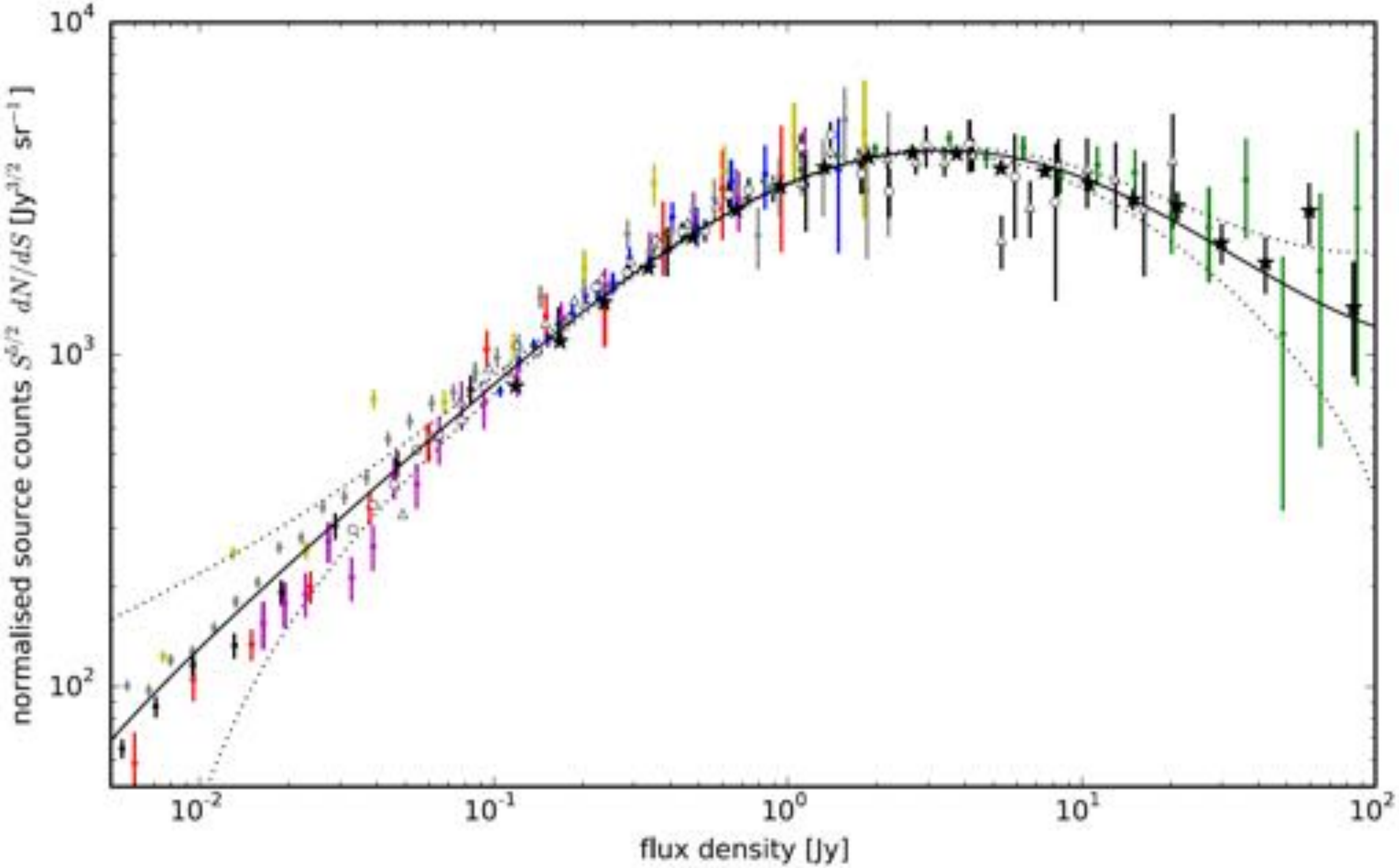
Resolution versus surface brightness sensitivity



TGSS as a reference survey

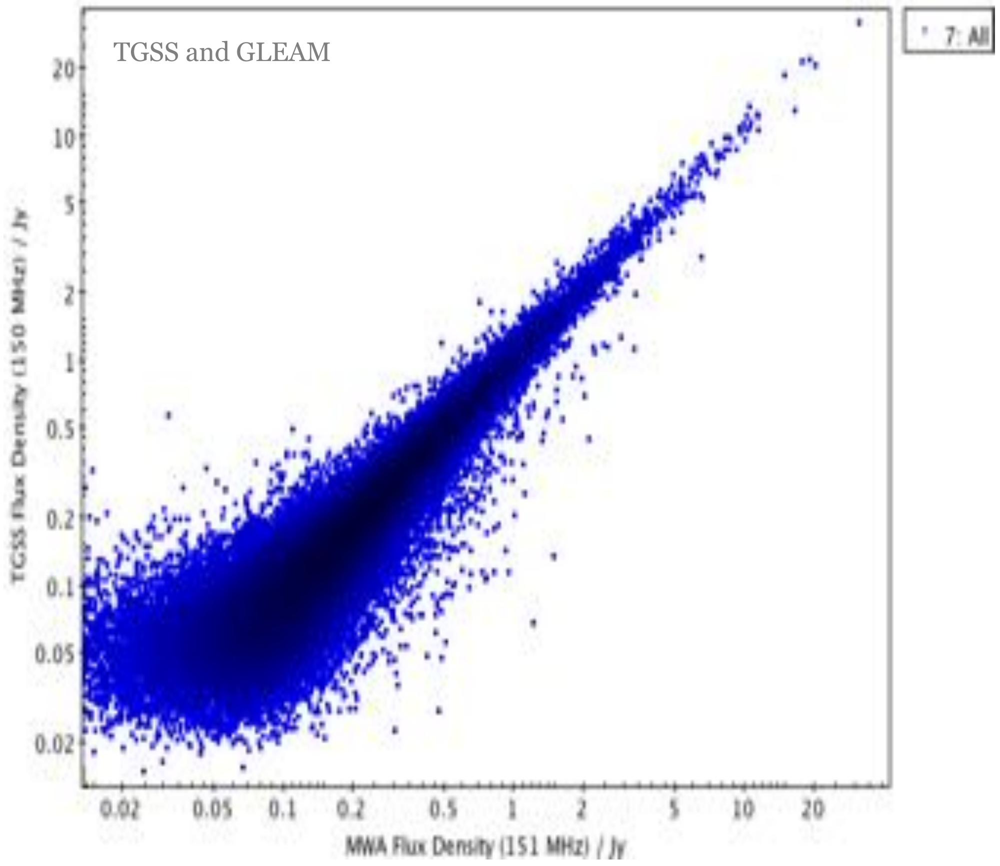
Detailed source counts at 150 MHz

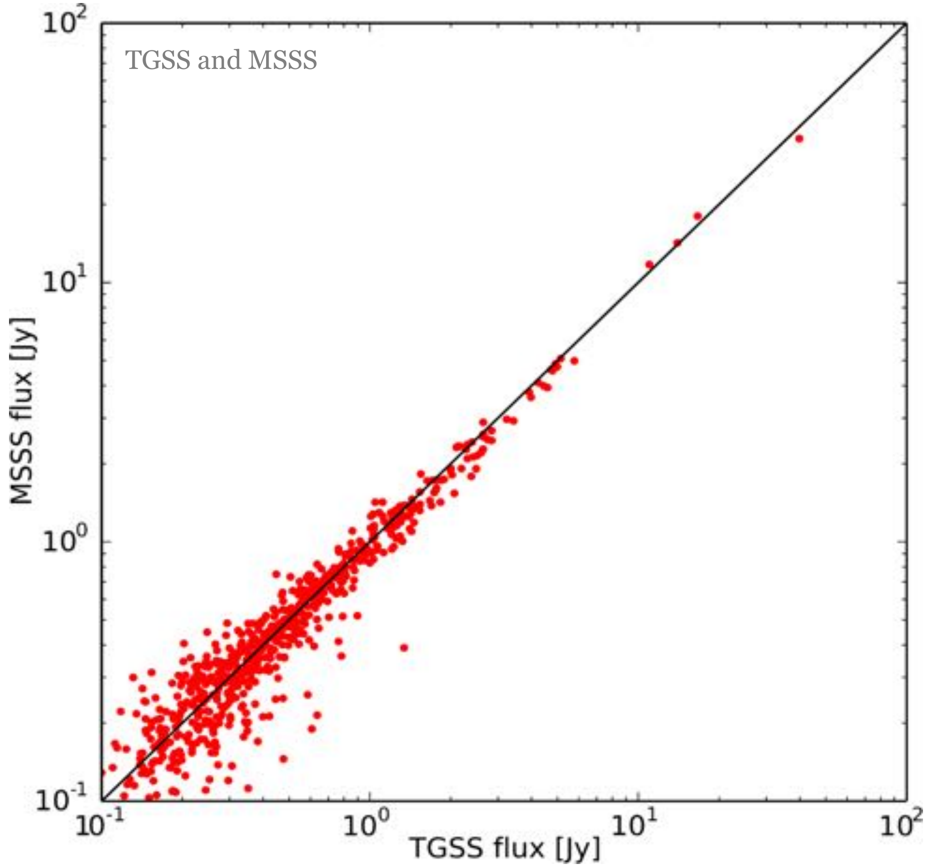
- Fitted with 6-parameter polynomial in log-log space
- See Intema+ 2016



TGSS as a reference survey

TGSS versus MWA-GLEAM and LOFAR MSSS flux density comparison



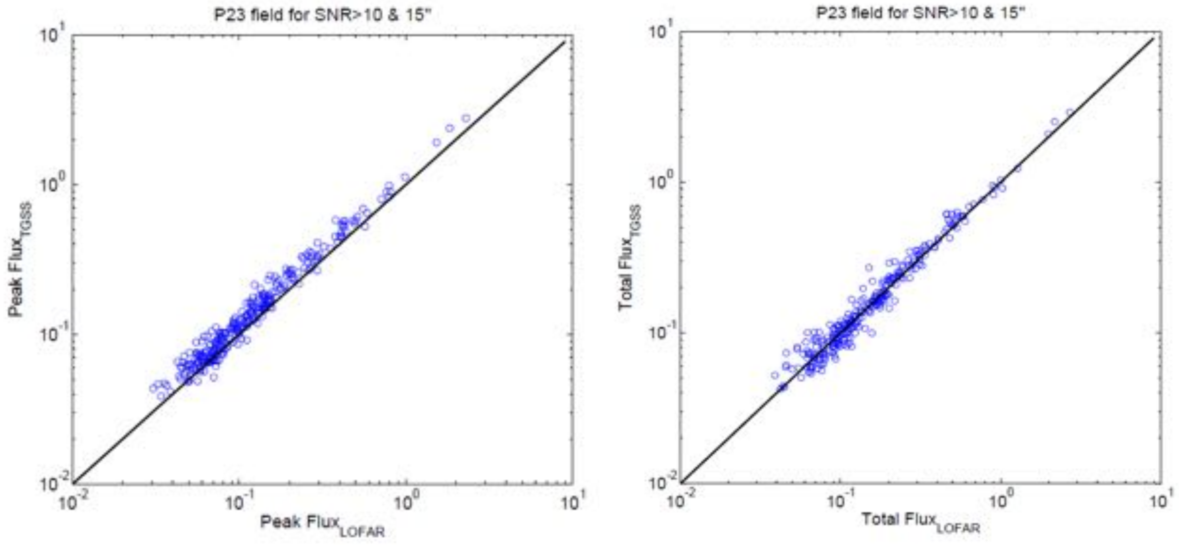


TGSS as a reference survey

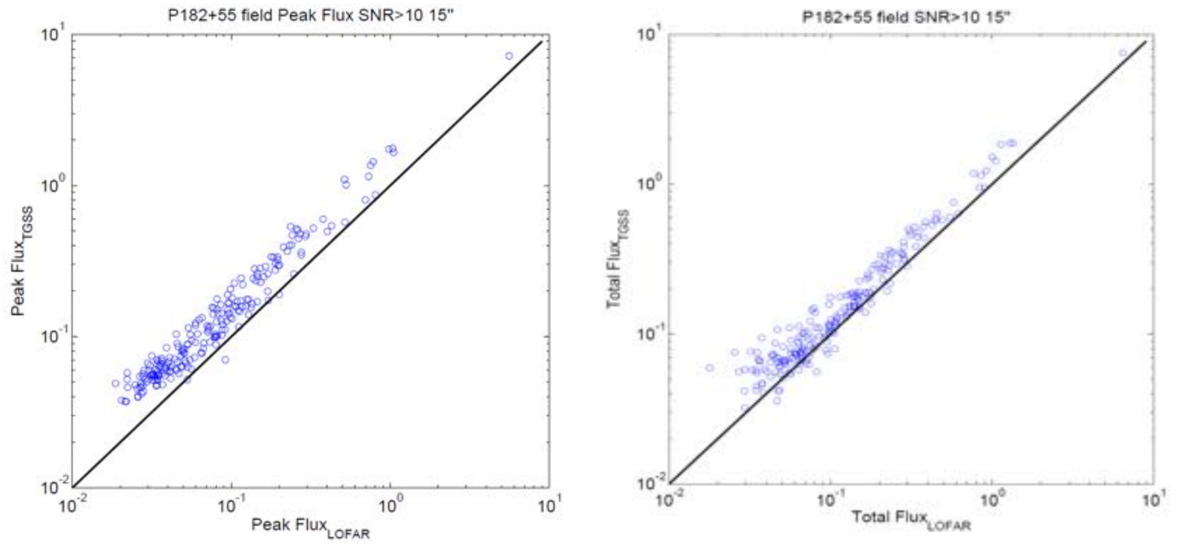
Quantifying ionospheric effects

- Mild ionospheric distortions causes angular broadening: Strehl ratio $R = \exp\left(-\frac{\sigma_\phi^2}{2}\right)$
- For unresolved radio sources: the ratio of peak flux over total flux
- Preliminary comparison between LOFAR HBA Tier-1 survey fields and TGSS survey data

mild ionospheric conditions

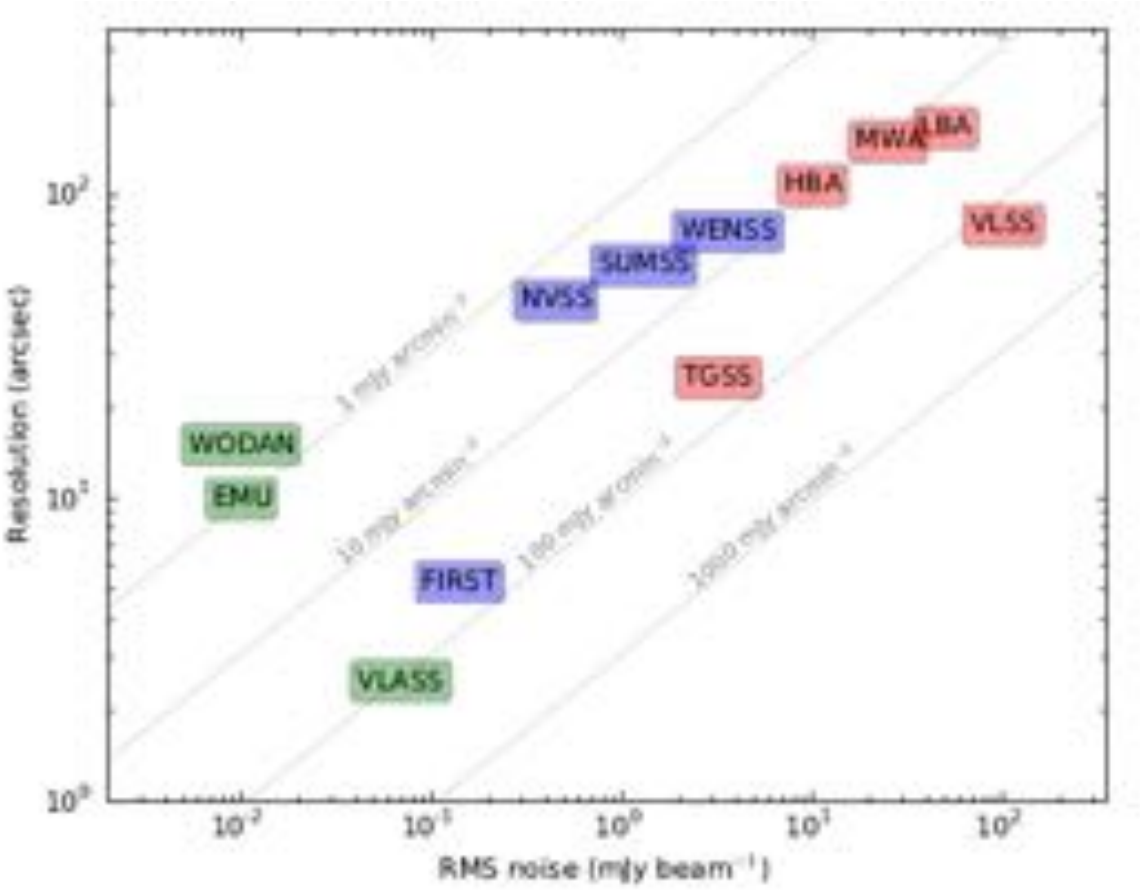


active ionospheric conditions



Current status and future plans

- TGSS ADR covers 90 percent of the radio sky at 25” resolution and a 2-5 mJy/beam median noise
- The full public data release includes
  - 5 x 5 deg² FITS images
  - Image cut-out service (up to 1 x 1 deg²)
  - Source catalog with 0.62 Million entries
- Resolution is better by factor of a few with respect to surveys at similar frequency
- Sensitivity is similar or better with respect to surveys at similar frequency
- Relatively high astrometric and flux density accuracy
- Estimated reliability is extremely high down to the 7-sigma detection threshold (>99 percent)



Current status and future plans

- TGSS images have been increasingly used as input models for LOFAR calibration
  - LOFAR Global Sky Model (GSM) generator now available on TGSS ADR webpage
- TGSS source catalog is used for cross-match by MWA GLEAM survey (to be released soon)
- TGSS data products are in itself interesting, mainly because of the large area covered
  - Cross-matched to NVSS to identify steep-spectrum radio sources (Mandal, de Gasperin, Emig)
  - Looking for radio counterparts of gamma ray sources detected by *Fermi* (e.g., Frail+ 2016)
  - Studying the spectral behavior of pulsars (compact steep-spectrum)
  - Finding high-z radio galaxy candidates based on their steep spectra (Saxena)
  - Targeted search for radio emission from exo-planets
  - Finding dying/dead radio galaxies
  - Finding merging galaxy clusters through their Mpc-scale diffuse radio emission
- http://tgssadr.strw.leidenuniv.nl/

Current status and future plans

Sterrewacht Leiden

- Impact of TGSS ADR1 is slowly increasing
  - 18 citations to main survey paper
  - ~10 visitors per day on project webpage
  - 100+ followers on Facebook page

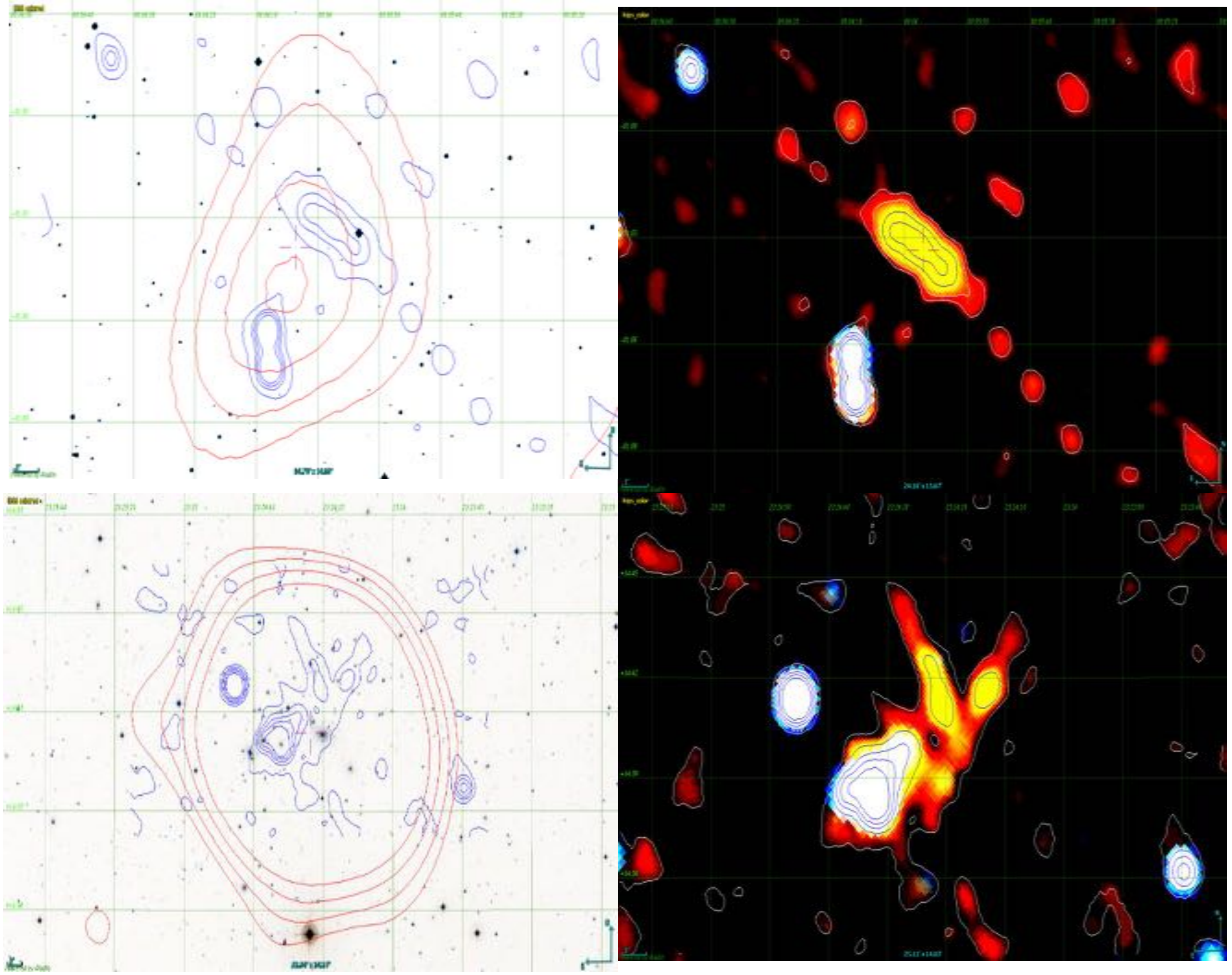
		Country
106 Hits	21.20%	Netherlands
99 Hits	19.80%	United States
95 Hits	19.00%	India
44 Hits	8.80%	Poland
37 Hits	7.40%	United Kingdom
30 Hits	6.00%	Australia
19 Hits	3.80%	New Zealand
13 Hits	2.60%	Germany
12 Hits	2.40%	
10 Hits	2.00%	France

- Various successful proposed follow-up observations
  - Candidate HzRGs with VLA, candidate milli-second pulsars with ATCA and GBT, candidate radio halos and relics in galaxy clusters (GMRT), candidate double-double radio galaxies (GMRT), ...
- Next data release (ADR2) planned for spring 2017
  - Incorporation of new make-up observations (70 hours allocated)
  - Incorporation of longer observations on extremely bright sources (A-team)
  - Possible inclusion of lowest declination fields (-53° to -55° DEC)
  - Flux scale corrections for subset of fields (~5 percent)

Science: A study of USSS in Galaxy Clusters

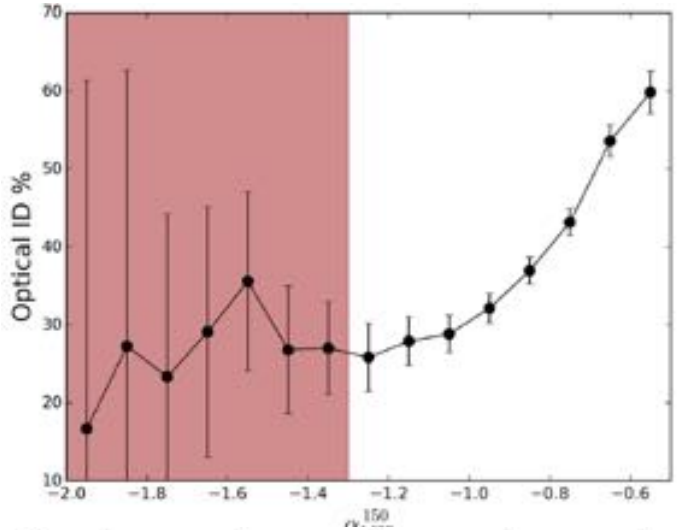
- Automated search for USSS in TGSS-NVSS with spectral index <~ -2
- Correlated to known cluster positions
- Subset of 7 sources selected for follow-up with GMRT at 325 MHz
- 35 hours allotted for GMRT cycle 31 (PI Soumyajit Mandal, PhD Leiden)

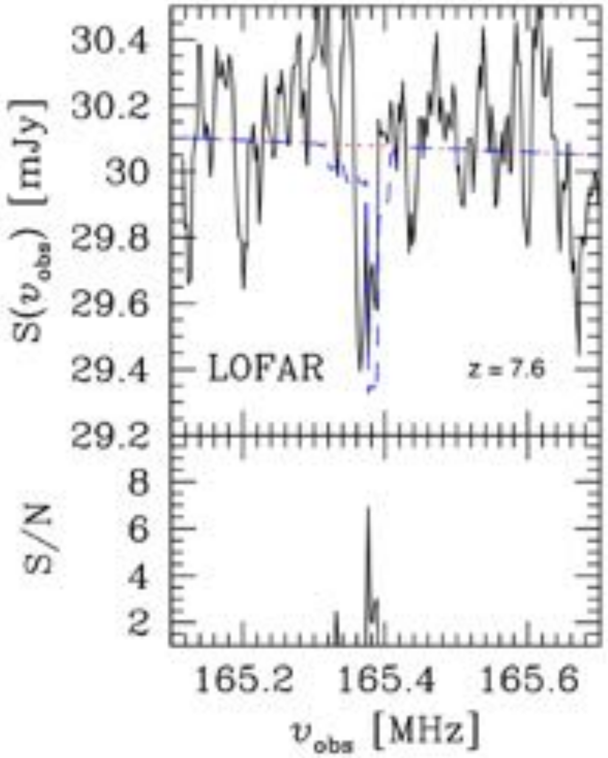
- Right: two USSS examples
  gray = DSS2R, red = ROSAT, blue = TGSS
  spectral index = -2.9 (top), -3.0 (bottom)

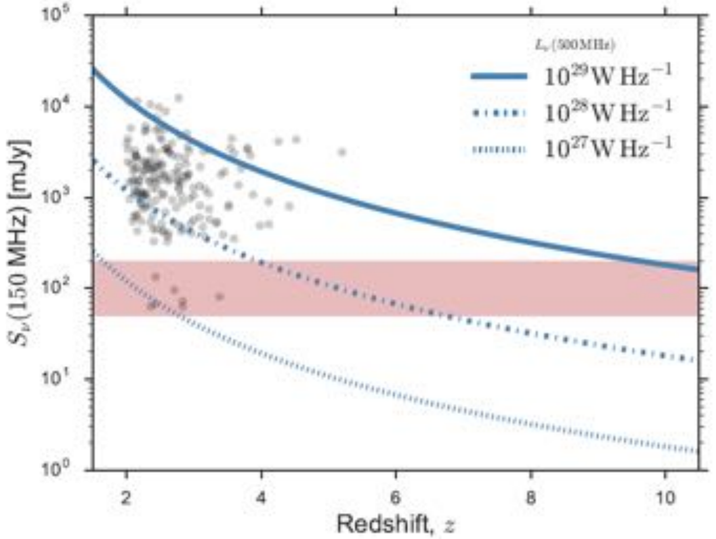


Science: High-Z Extreme Spectrum Project

- Selected compact USSS with TGSS-NVSS spectral index < -1.3
- Matched to FIRST to select on compactness and obtain good astrometry
- Ensure no optical and/or IR counterparts in all-sky optical/IR surveys (SDSS, PanSTARRs, WISE, UKIRT surveys)
- Selected 33 new USS sources for various follow-up (PI Aayush Saxena, PhD Leiden)
- Sample probes unexplored area in parameter space
- Hope to find the first HzRG at very high redshift (z>6) to probe HI absorption from the Epoch-of-Reionization







Science: Finding milli-second pulsars

- In TGSS steep spectrum PSRs (< -2.5) are (γ-ray emitting) MSPs (Frail+ 2016)
- Hope is to find exotic PSRs missed by traditional search methods
- Method selects without regard to period, DM, orbital parameters and scattering
- Measure compactness, spectrum, polarization position
- 11 PSR candidates in 3FGL, 12 in 4FGL
- VLA/ATCA follow-up for improved positions and resolution
- Promising candidates searched for pulsations in γ-ray (Fermi) and/or radio (GBT and Parkes)