# Processing of uGMRT data using the SPAM pipeline

Huib Intema

Senior Research Fellow

Curtin Institute of Radio Astronomy
International Centre for Radio Astronomy Research

# Scope

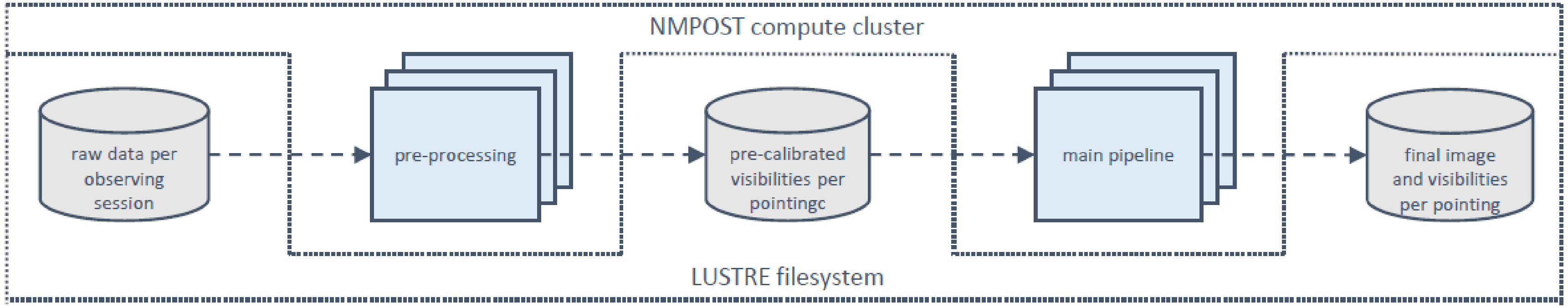
- Radio astronomy's large data era moves us towards automated data processing and multi-purpose large sky surveys
    - Personal interests: automated data processing, merging galaxy clusters, radio surveys
    - Heavy user of GMRT; projects already include many Indian collaborators
    - Relevant activities include:
        - SPAM algorithm for direction-dependent (ionospheric) calibration and imaging
        - SPAM end-to-end pipeline for automated processing of continuum GMRT and VLA data (< 1 GHz)
        - Alternative data release of the TIFR GMRT Sky Survey (TGSS) at 150 MHz

- Opportunities for new / improved collaborations
    - Increase user base of the SPAM pipeline; accelerate scientific throughput
    - Encourage further scientific exploration of the TGSS
    - Development of a pipeline for uGMRT, a dish-based pathfinder for SKA_LOW
    - Astrophysical explorations using (u)GMRT data in combination with complementary radio telescopes (MWA, ASKAP, ATCA, but also LOFAR, MeerKAT, JVLA)

# SPAM pipeline

- Robust, end-to-end data processing pipeline for GMRT (and VLA) low-frequency observations (Intema+ 2017)
  - Python/C implementation based on AIPS functionality, accessed through ParselTongue interface (Kettenis+ 2006) using Obit library (Cotton 2008)
- Performs wide-field (direction-dependent) ionospheric calibration, modeling, and imaging (Intema+ 2009) (SPAM = Source Peeling & Atmospheric Modeling)
- Publicly available, designed to deliver science-ready data products with minimal user interaction

- SPAM pipeline is split into a ‘pre-processing’ part and a ‘main pipeline’ part
- Each pipeline (part) instance runs as a single thread
- Parallelization only achieved when processing multiple data sets

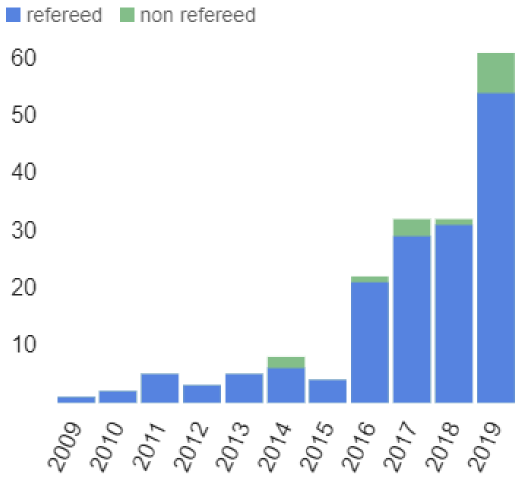



Data flow of the SPAM pipeline for processing the TGSS ADR1 project at NRAO
(from Intema+ 2017)

# SPAM pros and cons

- Successfully applied to 100's of GMRT observations, producing (near-) science-ready data products
  - Robust & fast processing, high success rate
  - Acceleration of scientific output
- Enabled very large projects like TGSS ADR1 and GAPP

- Dependence on 2013 AIPS and Obit poses some compatibility issues during installation
- Largely disconnected from latest algorithm developments
  - No multi-scale, multi-term deconvolution
  - No W-projection, W-stacking, etc.
- Limited to processing narrow-band stokes I observations with simple primary beam patterns
  - No A-projection, IDG, etc.



Publications referencing the SPAM pipeline (from ADS, 06/11/2019)

- Although some concepts are transferable, the SPAM pipeline itself has limited applicability
  - GMRT below 1 GHz(*), VLA at 4/P-band

- Developments for a wide-band SPAM based on CASA and/or LOFAR software have started (C-SPAM), but progress is slow

# SPAM pipeline for uGMRT?

- While uGMRT awaits the development of a full wide-band pipeline,
  the current SPAM pipeline can be used to process uGMRT frequency slices
  - Fractional bandwidth < 10-20 percent

- Functionality added to support wide-band data processing
  - Updated primary beam models
  - Updated processing parameters
  - Automated frequency slicing
    - 6x 33.3 MHz for uGMRT band 3 (300-500 MHz)
    - 8x 50 MHz for uGMRT band 4 (550-950 MHz)
  - Each frequency slice can be processed independently in parallel
    - Wide-band (GWB) processing time not too different from narrow-band (GSB)

# SPAM uGMRT band 4 example

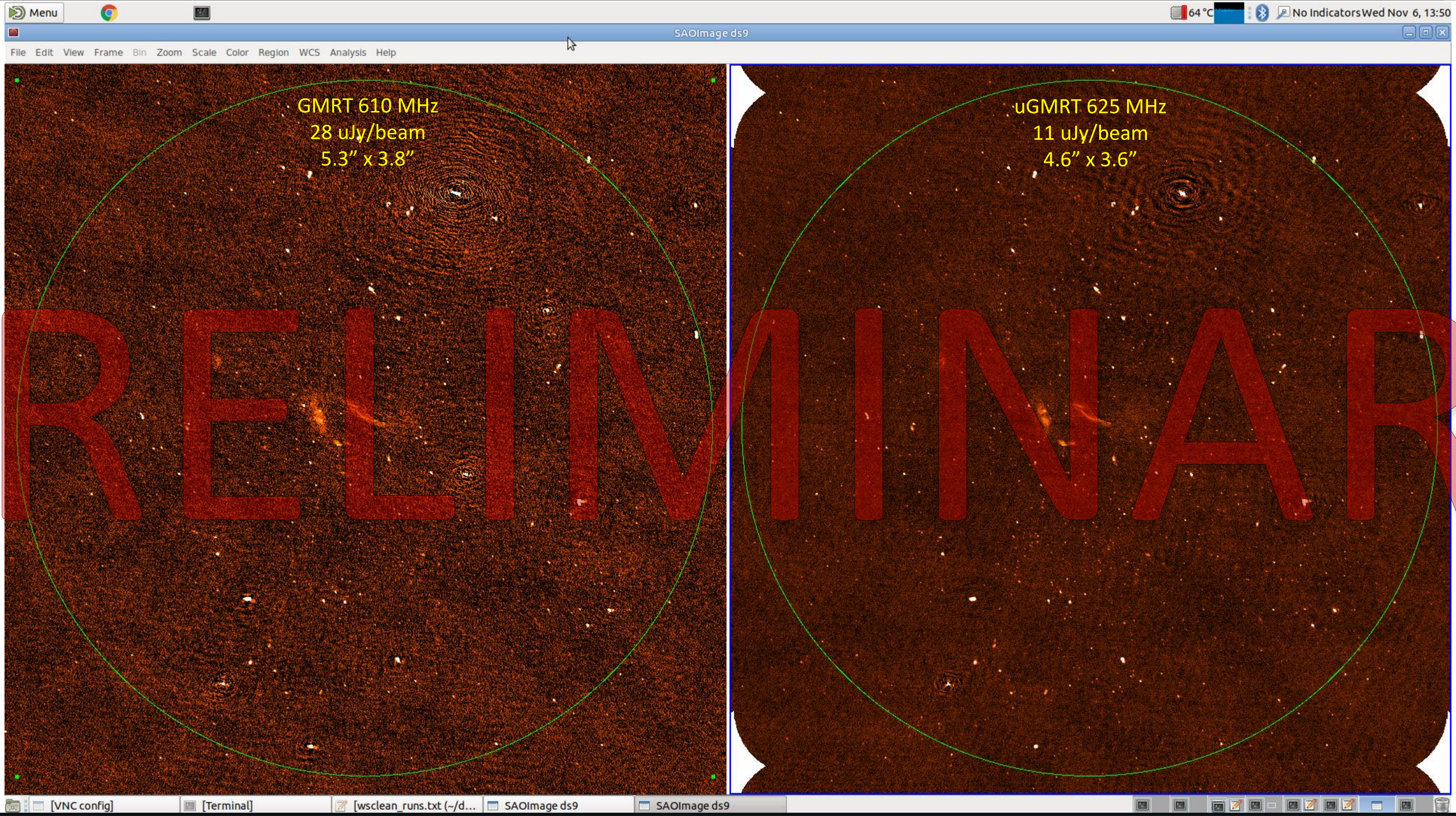
- uGMRT cycle 36 observations of merging galaxy cluster
  - 8 hours in band 4 (550-750 MHz) yielded 5.9 hours on-target
  - 610 MHz narrow-band data (GSB: 33.3 MHz) recorded alongside wide-band data (GWB: 200 MHz)

- Please respect that these are preliminary results on proprietary data, presented with permission of the PI

- Band 4 data processed in 4x 50 MHz frequency slices
  610 MHz narrow-band data (33.3 MHz) processed in parallel
- SPAM data processing done within 2.5 days
  - Produced final image and calibrated visibilities

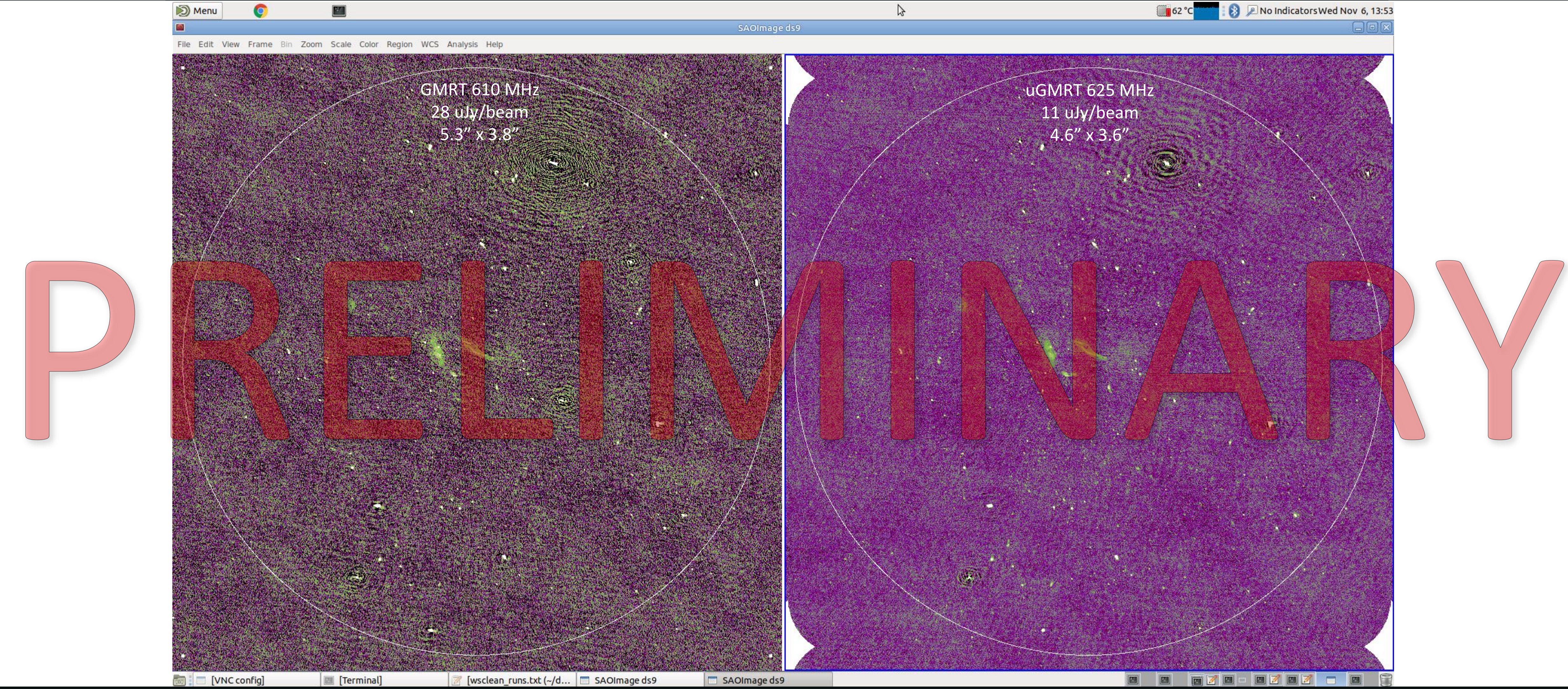
- Following images are not corrected for primary beam attenuation …

# SPAM uGMRT band 4 example (2)

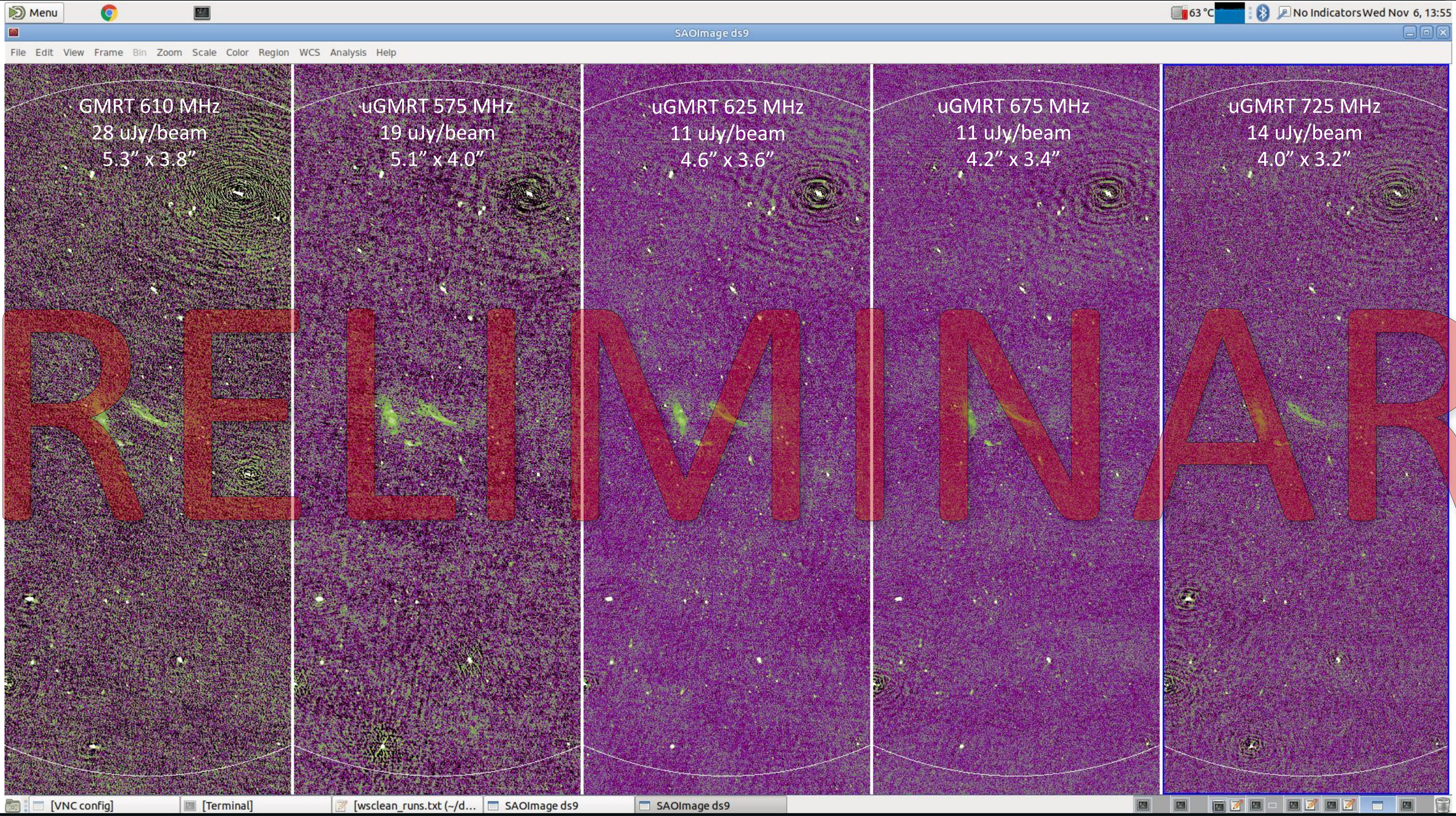



PRELIMINARY

# SPAM uGMRT band 4 example (3)



# SPAM uGMRT band 4 example (4)

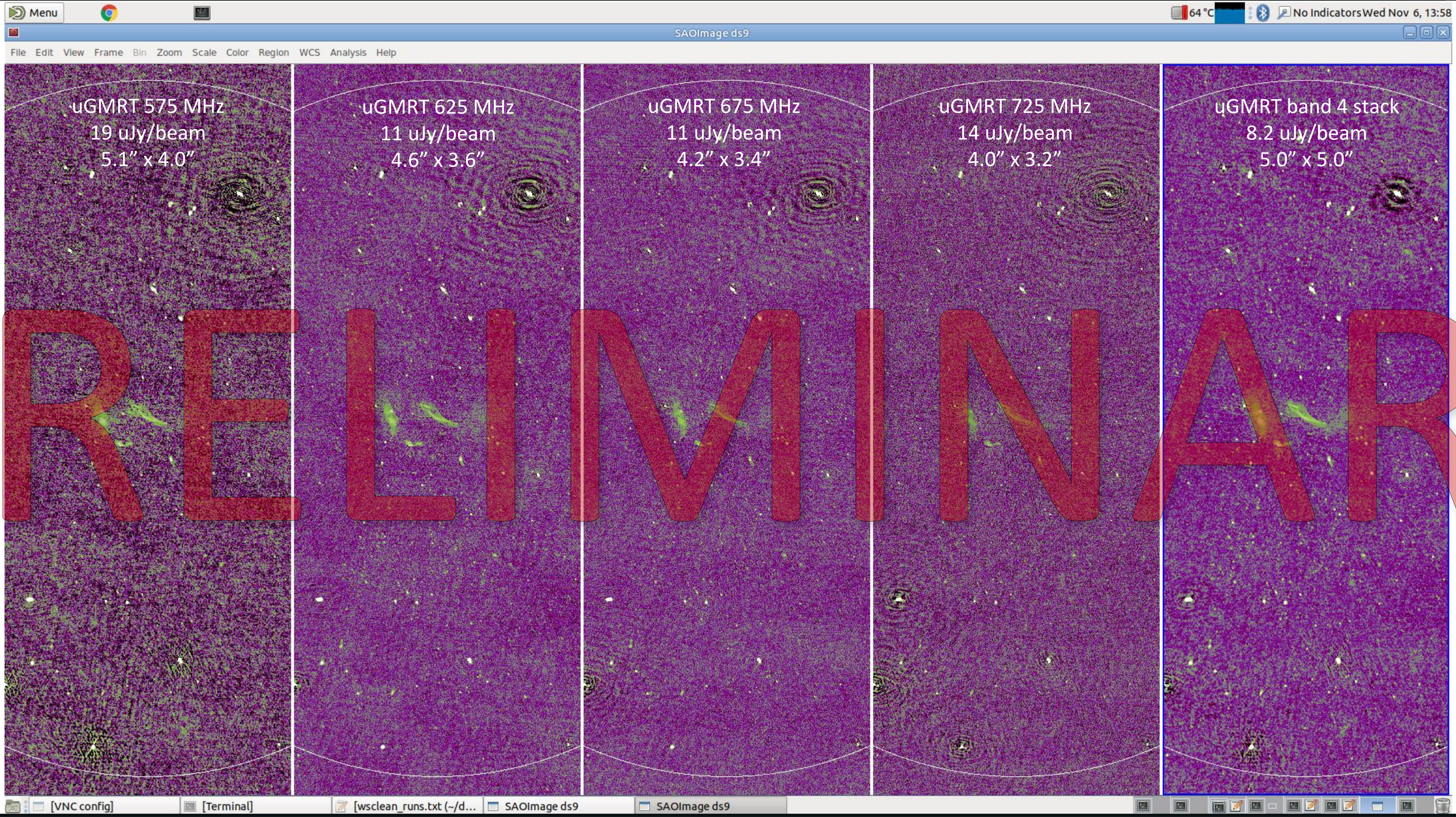


# SPAM uGMRT band 4 combination

- uGMRT band 4 performance is superior compared to GMRT 610 MHz
  - Much better than naïve bandwidth scaling (33.3->50 MHz gives factor 1.22)
- Combining the 4 frequency slices should yield even better results
- Two possible approaches:
  - Image plane combination (linear stacking)
  - UV plane combination (concatenating visibilities)

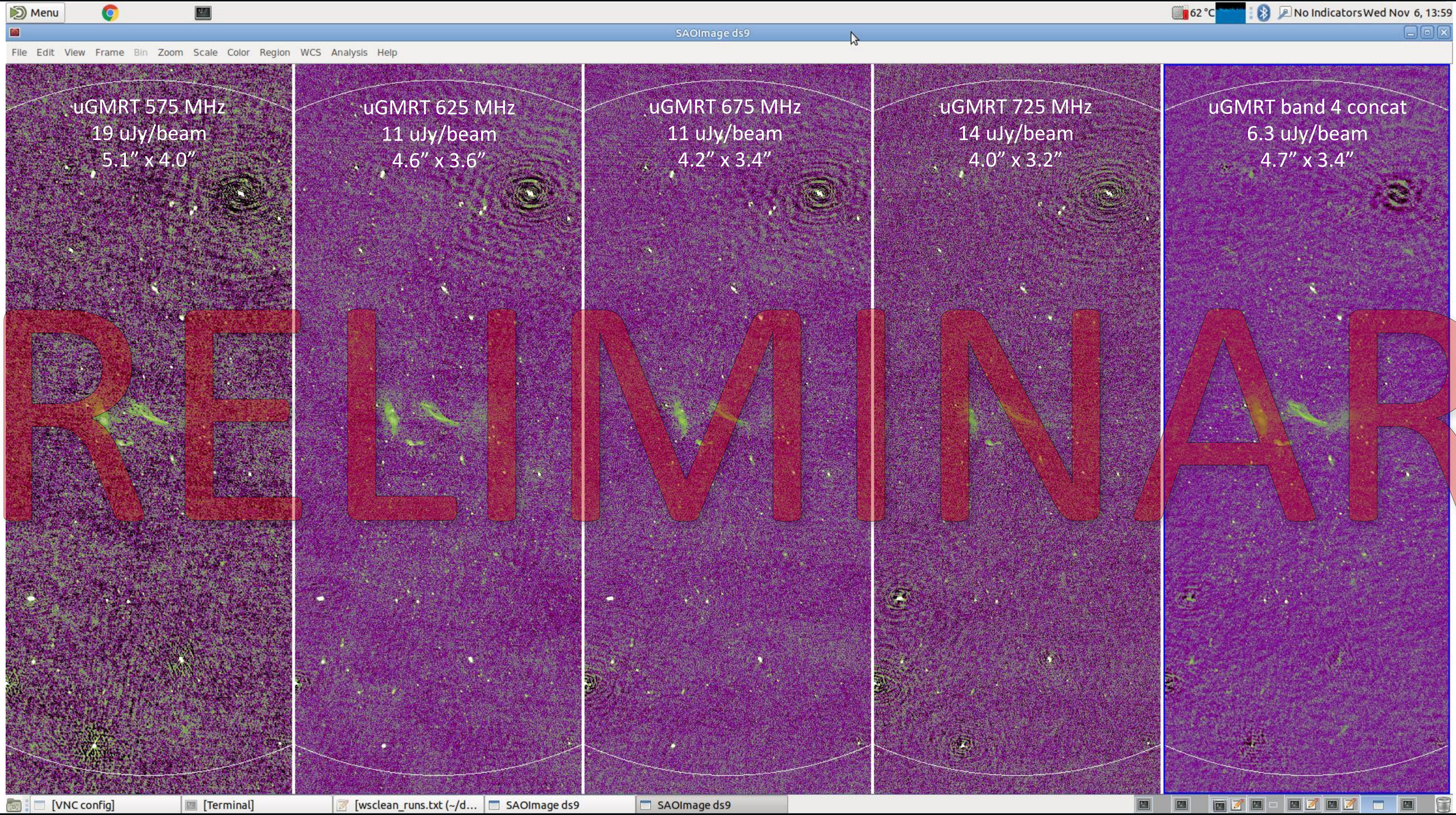
- Image plane combination:
  - Simple, quick operation
  - Robust against small flux and astrometric offsets between images
  - Limited by resolution of lowest frequency slice
  - Low-level artifacts due to CLEAN limit in images of individual frequency slices
- UV plane combination:
  - Imaging at full resolution possible
  - Deeper CLEANing than in images of individual frequency slices
  - Requires concatenation and re-imaging of visibilities
  - Small flux and astrometric offsets will introduce additional image artifacts

# SPAM uGMRT band 4 combination (2)

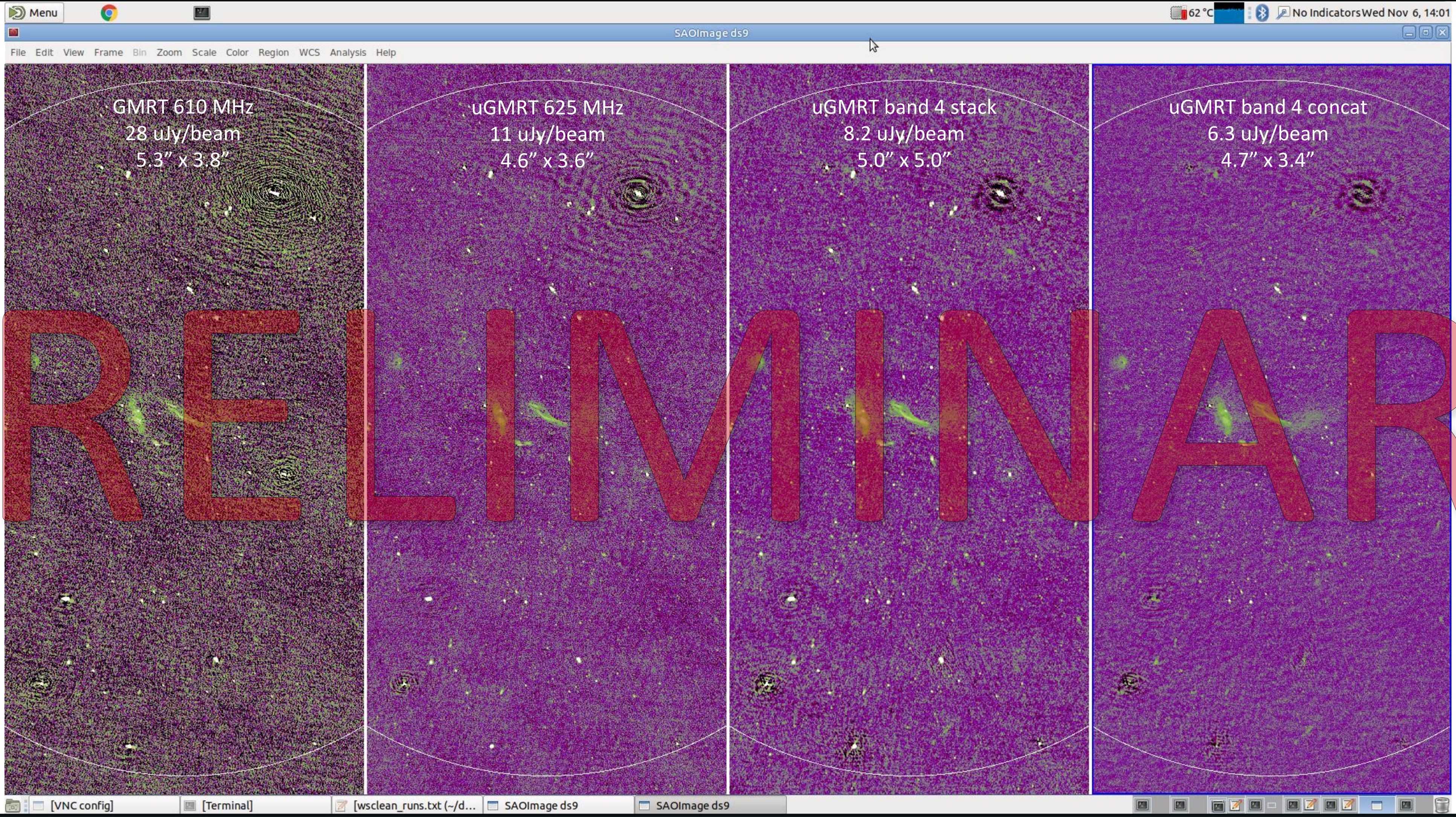



PRELIMINARY

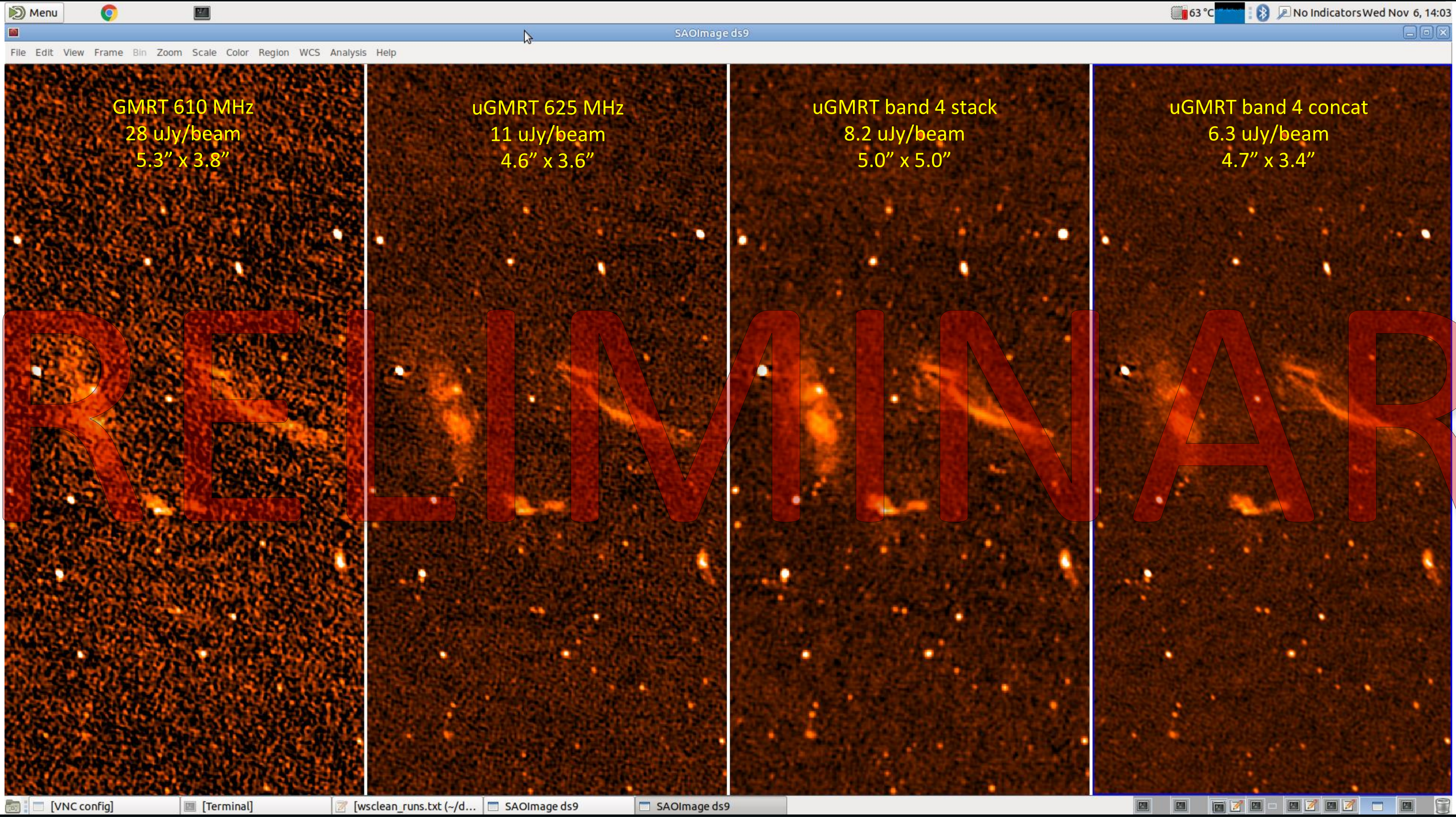
# SPAM uGMRT band 4 combination (3)



# SPAM uGMRT band 4 combination (4)



# SPAM uGMRT band 4 combination (5)




PRELIMINARY

# SPAM uGMRT summary

- UV plane combination yields the best sensitivity at high resolution
  - In this example, 4-5 times better sensitivity than the legacy 610 MHz band
  - Note that no primary beam corrections were applied; needs to be incorporated
  - Experiments with uGMRT band 3 shows similar behavior

- UV plane combination requires
  - An imager that supports joined deconvolution of independently flux-calibrated frequency slices (WSClean)
  - Careful preparations when processing the frequency slices
    - Matching frequency bands and resolution
    - Accurate astrometric alignment

- Approach still needs a way to correct for frequency-dependent primary beam attenuation
- Automated generation of an in-band spectral index map will be added value

- Full recipe will be published on the SPAM webpage

# Radio-optical image


PRELIMINARY