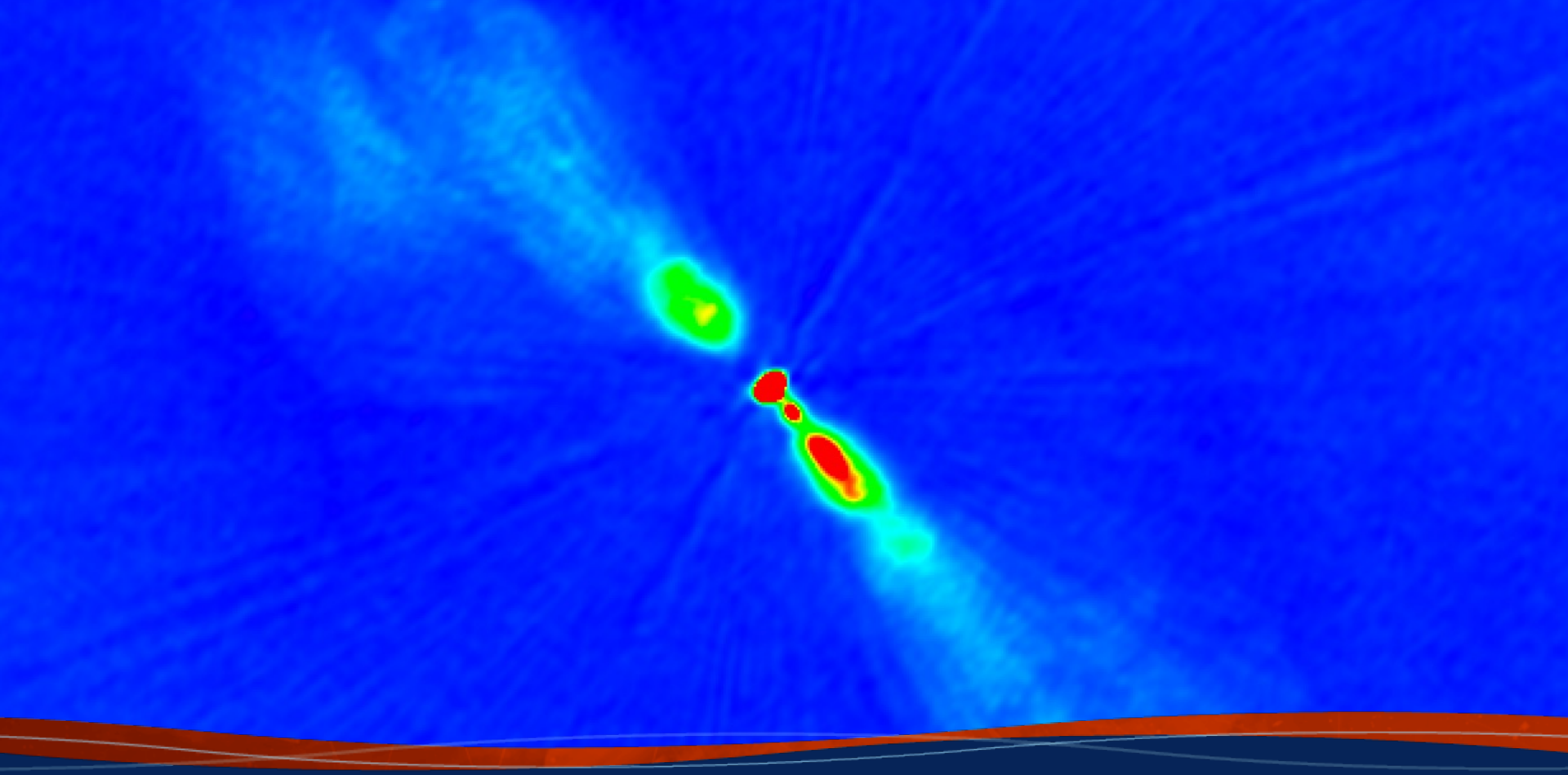


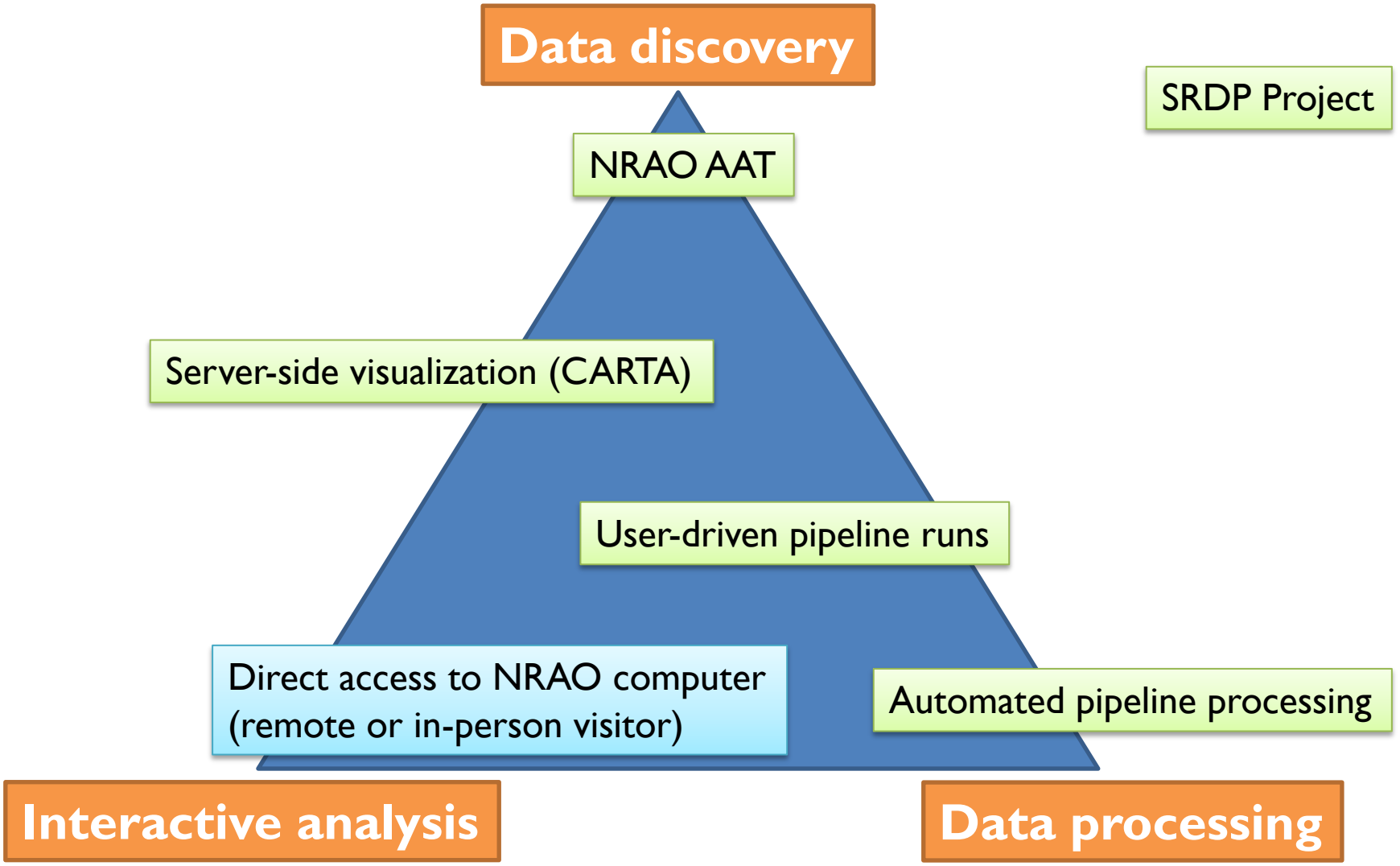
VLA Science Ready Data Products

Mark Lacy

Outline

- VLA Science Ready Data Products (SRDP) for PI-driven projects
  - SRDP as a Science Platform for NRAO
  - What do we mean by "science ready"?
  - Current operations
    - Calibration of C through Q band continuum
    - Imaging pilot
    - Image display over the web using CARTA (John Tobin's talk)
  - Near-Future scope (~next 4-12 months)
    - Imaging with self-calibration
    - Gain compression fix, L and S-band continuum calibration
    - Large project archiving
  - Longer-term scope (~1-3 years)
    - User-driven calibration and imaging
    - Imaging from combined EBs
    - Spectral line calibration and imaging
- VLA Sky Survey (VLASS) and VLASS data products
- What SRDP has taught us

SRDP as a science platform for NRAO



Courtesy E. Ryan

What do we mean by "Science Ready"?

- We define science ready products as being processed and quality assured to a uniform standard, with known provenance - ``publishable as is", but also a basis for more advanced processing.
  - For simple Stokes I continuum science cases, VLA SRDP calibration will be good to go.
    - Spectral line and/or polarization calibration are not yet supported, see later.
  - The standard imaging pipeline imaging uses a fixed robust=0.5 and happens only on a per-EB basis. Combining data from different EBs is left to the user. Self-calibration is not currently done, but is coming soon.
    - Images should be useful as a reference, many users will want to reimage.
    - User-driven imaging, with options for tuning beamsize and combining EBs from either a single or multiple configurations is planned.

Current SRDP Operations for the VLA

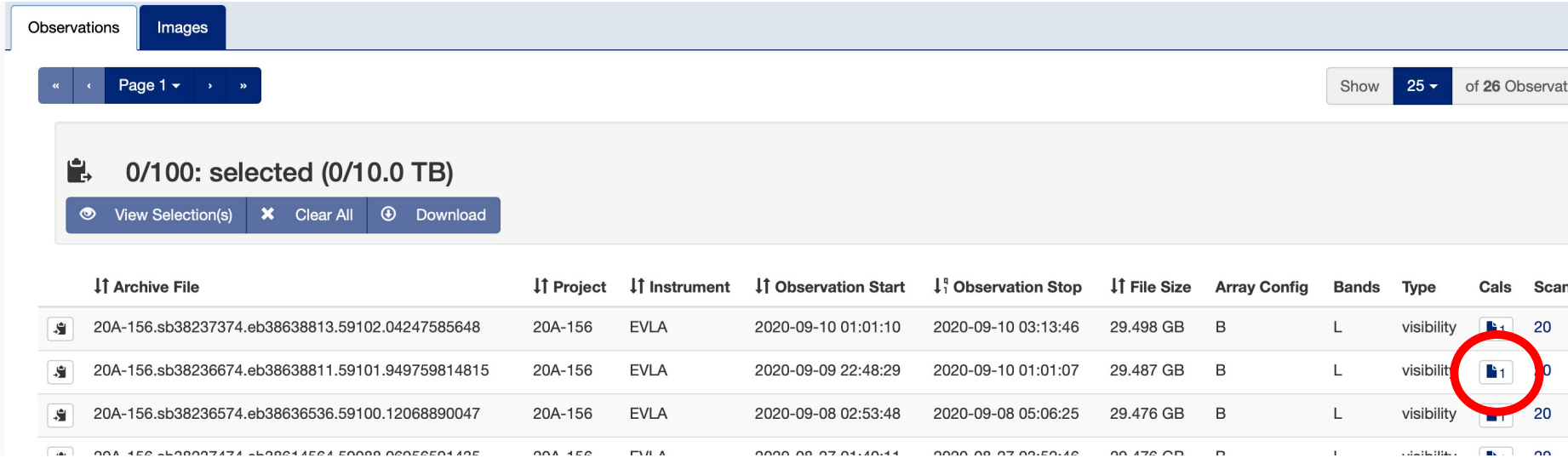
- Qualifying for SRDP is decided on a per-SB basis:
  - C, X, Ku, K, Ka or Q-band (L and S-band to be added in a few months)
  - Wide bandwidth useful for continuum science (>~ 1GHz) *(pipeline is not optimized for spectral line science, see https://science.nrao.edu/facilities/vla/data-processing/pipeline#section-26 for suggestions on running the pipeline for spectral line data)*
  - SBs should have a standard calibration strategy, and use standard intents.
  - Phase calibrators should not be significantly resolved.
  - Observations should use either 3C286 or 3C147 as flux calibrators (3C48 and 3C138 have varied recently, and updated/time-dependent models are not yet available in CASA/pipeline).
- Imaging is at present a pilot program, but will expand over the summer.
- EBs that do not qualify for SRDP are still sent through the calibration pipeline, but the results are not checked in detail and should not be considered ``science ready".

QA review and flagging

- Data that qualify for SRDP are reviewed in detail by a data analyst (DA).
- Extra flags are applied for bad data missed by the pipeline.
  - Typically minor issues related to RFI missed by the pipeline, poor pointing (at high frequencies), weather, data transfer problems, low-level (<20%) gain compression on a few antennas.
- Data are rerun with the extra flagging.
- Flagging is iterated if necessary and data rerun
- Results are inspected by an Astronomer on Duty (scientist or senior DA) before being archived.
- Image QA is a separate step.
  - Images can be rerun with e.g. extra flagging or different clean parameters.
  - Results checked by an AOD for artifacts etc before archiving.

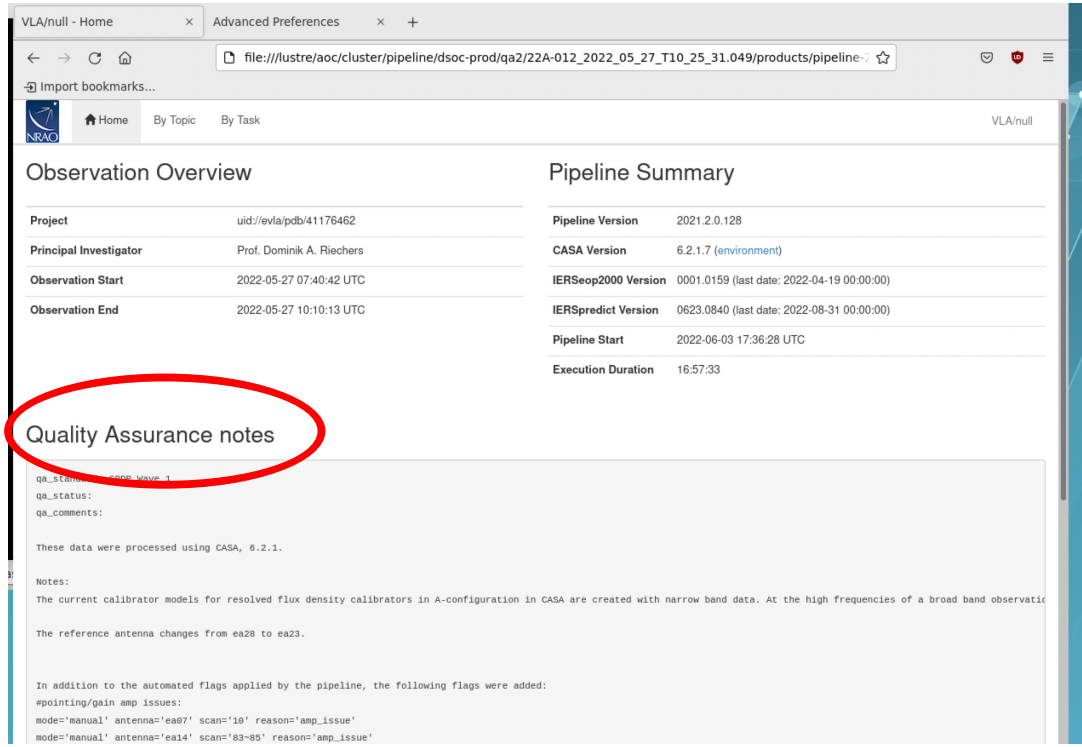
Navigating your calibration weblog

- Detailed information on the calibration of your EB can be found in the weblog that is produced by the pipeline.
- Go to the archive, observations tab and click on the "cals" icon to retrieve the tarred up calibration files and weblog.



Navigating your weblog

- The VLA CASA Pipeline Guide (casaguides.nrao.edu/index.php/VLA_CASA_Pipeline) contains a detailed description of the weblog.
  - You may also need to change the security settings in your browser to view some of the pages as described in the Guide.
- First, look at the index page of the weblog, where QA notes from the DAs appear. Another useful page is the Plot Summary (stage 17) where a number of diagnostic plots are presented.



Plot Summary

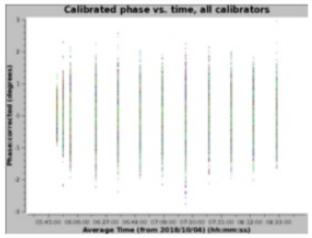
17. Plot Summary

BACK

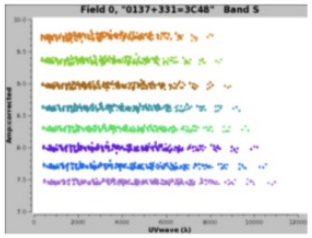
VLA Pipeline Summary Plots

Phase vs. time for all calibrators, Amp vs. UVwave for all calibrators, as well as a representative selection of fields with intent='TARGET' with Amp vs. UVwave plots.

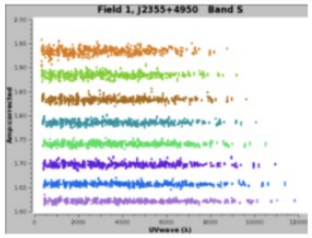
TDRW0001.sb35624494.eb35628826.58395.23719237269.ms



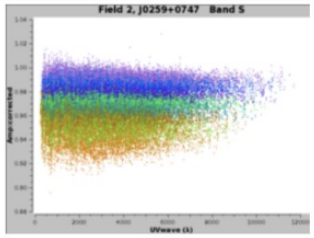
All calibrators Band: All bands



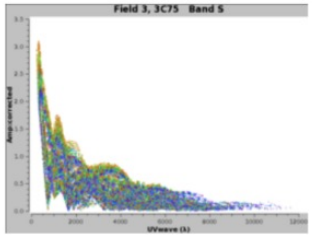
Field 0, "0137+331=3C48" Band: S



Field 1, J2355+4950 Band: S



Field 2, J0259+0747 Band: S

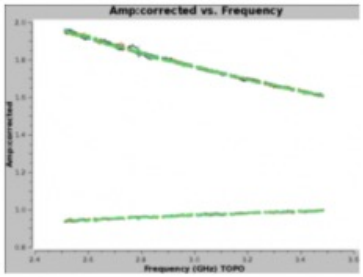


Field 3, 3C75 Band: S

Calibrated amplitude vs frequency

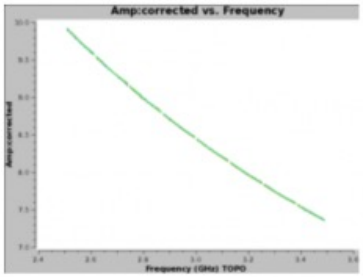
Plots of calibrated amplitude vs frequency for all antennas and correlations, coloured by antenna.

TDRW0001.sb35624494.eb35628826.58395.23719237269.ms



Receiver bands: 13cm (S)
(Spws 2, 3, 4, 5, 6, 7, 8 and 9)

Phase calibrator: J2355+4950 and J0259+0747.



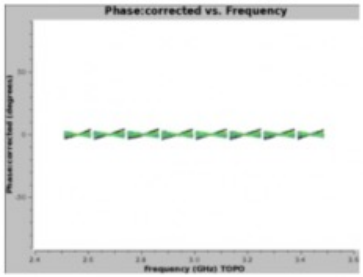
Receiver bands: 13cm (S)
(Spws 2, 3, 4, 5, 6, 7, 8 and 9)

Bandpass calibrator: "0137+331=3C48".

Calibrated phase vs frequency

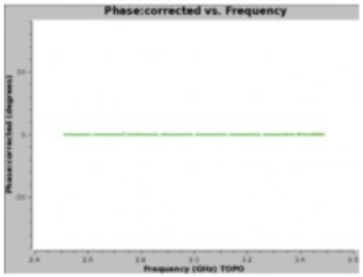
Plots of calibrated phase vs frequency for all antennas and correlations, coloured by antenna.

TDRW0001.sb35624494.eb35628826.58395.23719237269.ms



Receiver bands: 13cm (S)
(Spws 2, 3, 4, 5, 6, 7, 8 and 9)

Phase calibrator: J2355+4950 and J0259+0747.



Receiver bands: 13cm (S)
(Spws 2, 3, 4, 5, 6, 7, 8 and 9)

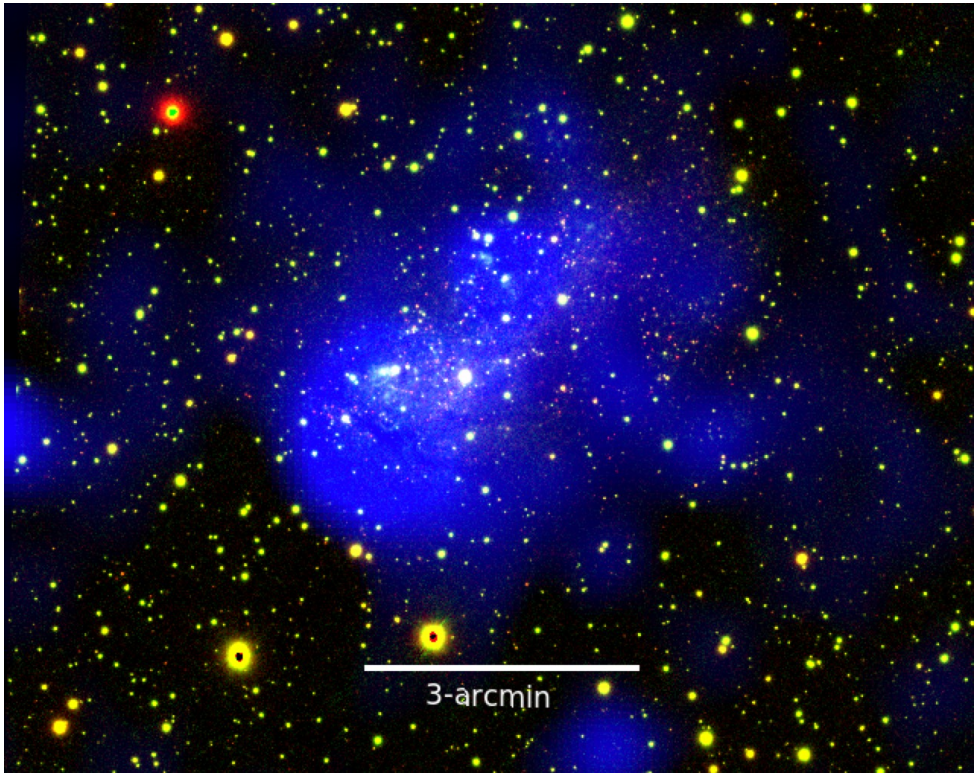
Bandpass calibrator: "0137+331=3C48".

Navigating the weblog

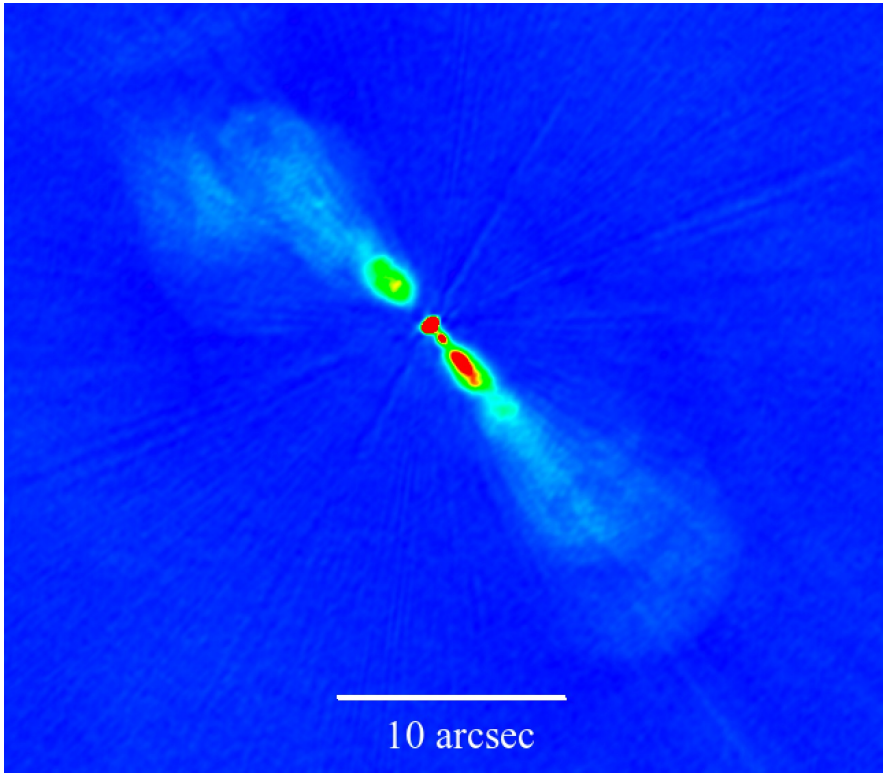
- There is also a weblog for imaging, though these tend to be less useful. Currently they are not retrievable from the archive, but can be obtained on request.
- From this August, the initial calibration weblog (before reruns by the DAs) will also be made available for processed data, so users can compare the data before and after manual flagging.
- If, after examining the weblog, you have questions about flagging etc, please feel free to contact the helpdesk.

Example pipeline images

IC10 (L-band, D-config)



Field FRI C-band, A-config



Why did my EB fail SRDP calibration?

- Failure reasons include
  - >~20% gain compression on multiple antennas (affects flux density scale and calibration quality).
  - Resolved phase calibrator (>~10% resolved)
  - Poor weather with rapidly changing conditions (at high frequencies).
  - Scan intents incorrect (rare)
  - OPT bugs or user errors in scheduling resulting in severe shadowing or complete loss of a calibrator observation (very rare).
  - Flux calibrator variability (very rare).
  - ~<18 working/unflagged antennas (e.g. some subarrays, or move-time obs) (rare).
- An SRDP fail can usually be recovered by manual reduction.
  - Note that a gain compression fix (not yet available in CASA) will be available in the next version of the pipeline, to be released this Fall.

Operations metrics

- Over the course of SRDP (since June 2019) the median time from observation to delivery of the calibration has been 7 days.
  - This does depend a lot on the configuration, A-configuration, for example, is slower (10-20 days) due to the higher data volume. Also holiday breaks slow us down, as do configuration moves (we need to wait for the antenna positions to be measured before running the pipeline).
  - Imaging will add another 3-5 days, but the calibration products are available to the proposers as soon as they are ingested, before the imaging pipeline run starts.
- The average number of reruns (after the initial pipeline run) is1.6.
  - Some datasets need no reruns, some 4-5, more typical is 1-2.

Pause for questions on current SRDP Ops

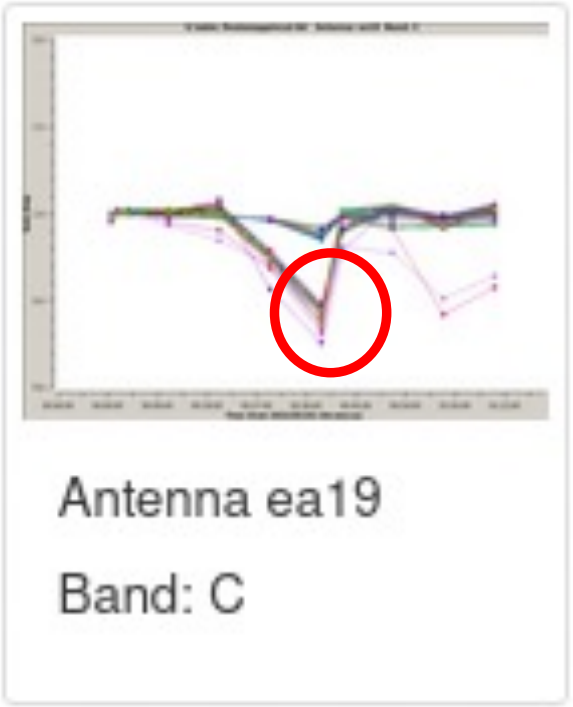
Near-future (next 3-4 months)

- Gain compression
- Total Electron Content (TEC) maps
- Self calibration of images

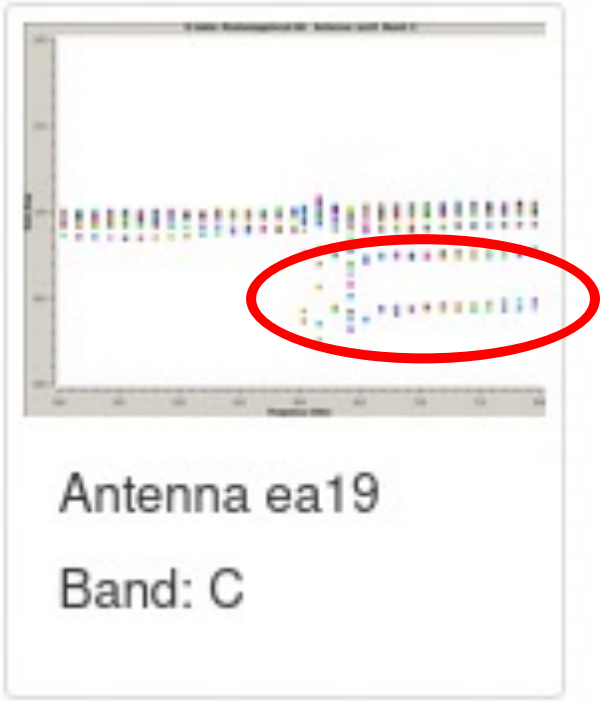
Gain compression

- Happens when severe RFI is detected to the extent that it changes the apparent brightness of the sky as seen by the telescope and thus changes the system gains.
  - Effect is per baseband, so most of baseband can be unaffected by RFI, but still get affected by gain compression.
  - Recognize by seeing calibration amplitude variations that affect an entire baseband.
  - Can also be seen in the switched power plots (hifv_priorcals)
  - Mostly affects L- and S-bands, but now being seen at higher frequencies (especially C-band)
- Fix is to use the switched power data to correct for the change in amplifier gains.
  - Will be in the pipeline from ~October 2022
  - Correction is best for 8-bit (so if you want high flux density accuracy best to use 8-bit), but seems to work reasonably well for 3-bit data too.

Gain compression from hifv_finalcals plots (stage 12)

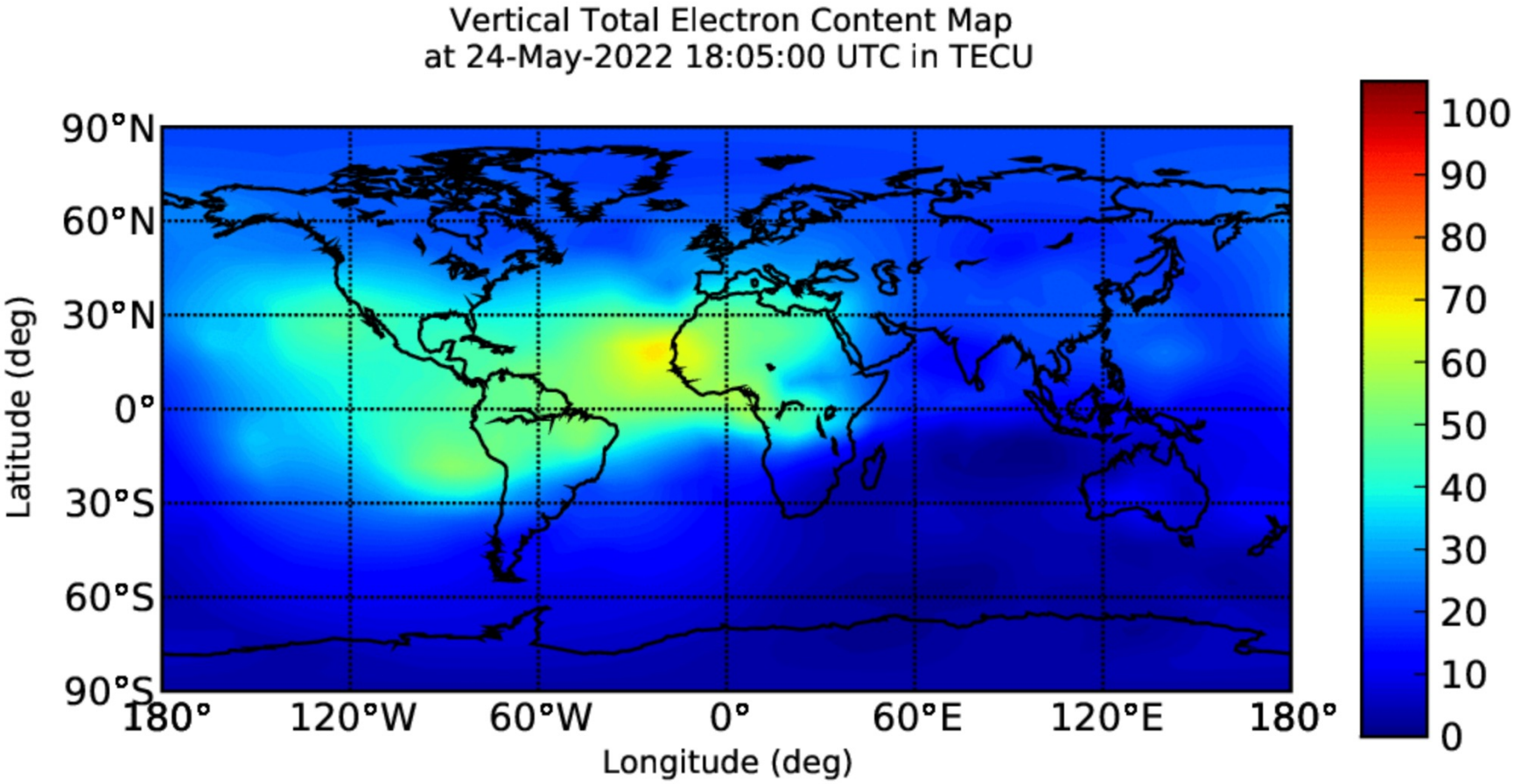


Amp correction vs time



Amp correction vs freq

TEC maps



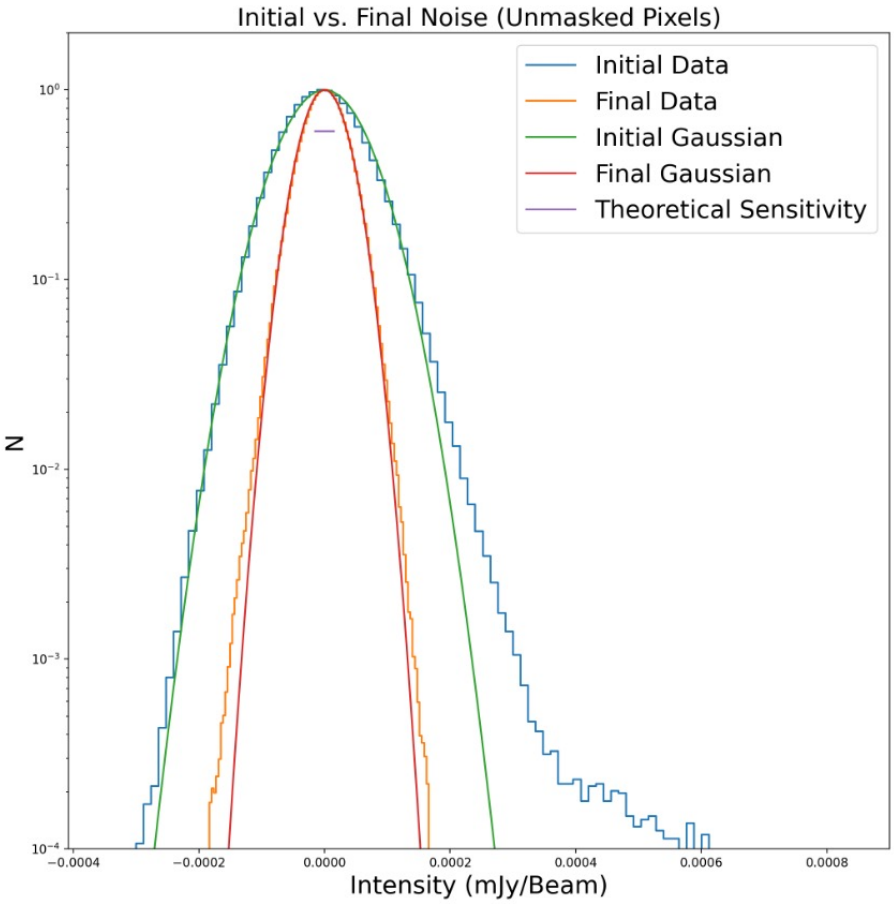
- Using dispersion of the signal from GPS satellites, NASA makes maps of the total electron content of the ionosphere with a cadence of once every 2hr,

TEC maps

- The VLA pipeline uses these to estimate the phase delay, dispersion and Faraday Rotation through the ionosphere.
- The maps from NASA lag by about 2 weeks, so L-band calibrations will be delayed to allow the TEC maps to be used in cases where polarimetric data are taken. In the long term, the application of the TEC maps will be enabled through user-driven calibration runs.

Imaging: Self-calibration

- Self-calibration takes advantage of the fact that the complex gain solutions from any given antenna are over-determined (because many baselines are sampled by an interferometer for a given antenna.).
- The SRDP imaging pipeline will self-calibrate in phase only, iteratively choosing a solution interval to obtain the highest dynamic range image.



Tobin, Yoon et al. in prep

Far-future (next 1-3 years)

- Large projects/collections
- SRAO pipeline interface
  - Spectral line calibration and imaging
  - User-driven calibration and imaging
    - Imaging from combined EBs
    - Multi-configuration imaging
  - Polarization calibration and imaging
- Large project/collection support
- *[TTA tools (new proposal/scheduling software)]*

Large projects/collections

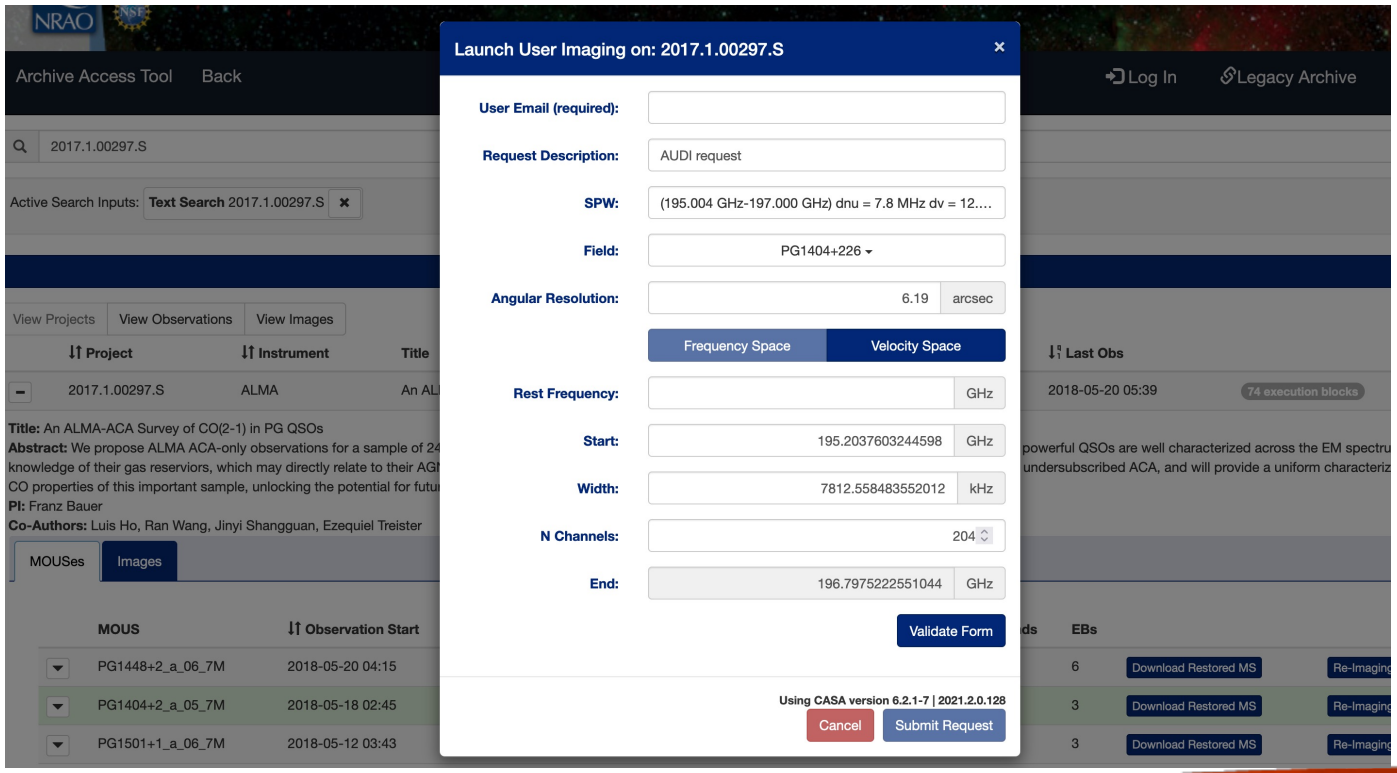
- Users with large projects on the VLA and other radio telescopes will soon be able to have the final data products hosted in the NRAO archive as ``collections''.
- Collections will have their own web pages, but also be searchable in the general archive interface.
- Digital Object Identifiers (DOIs) will be assigned to each collection to facilitate data publication.

Spectral line calibration and imaging

- The current VLA pipeline is optimized for continuum
- To adapt the calibration pipeline for spectral lines a couple of changes are needed:
  - Turning off (or reducing) edge channel flagging for spectral windows.
  - Algorithms for dealing with calibrations in narrow channels and spectral windows (e.g. bandpass smoothing).
  - Supplying a cont.dat file that defines which regions contain only continuum. Regions not included in cont.dat are assumed to contain lines, and therefore RFI flagging is not applied to them.
    - Prior knowledge of the line frequencies is therefore needed. This will be provided by user input through the TTA tools package.
- The spectral line imaging pipeline will be similar to that for ALMA.

User-driven VLA calibration and imaging

- In some cases, a user may wish to recalibrate or reimage data. Reimaging is likely to be the most common use case, and, indeed for ALMA we have the ALMA User-Driven Imaging (AUDI) interface working already.



User-driven VLA recalibration and imaging

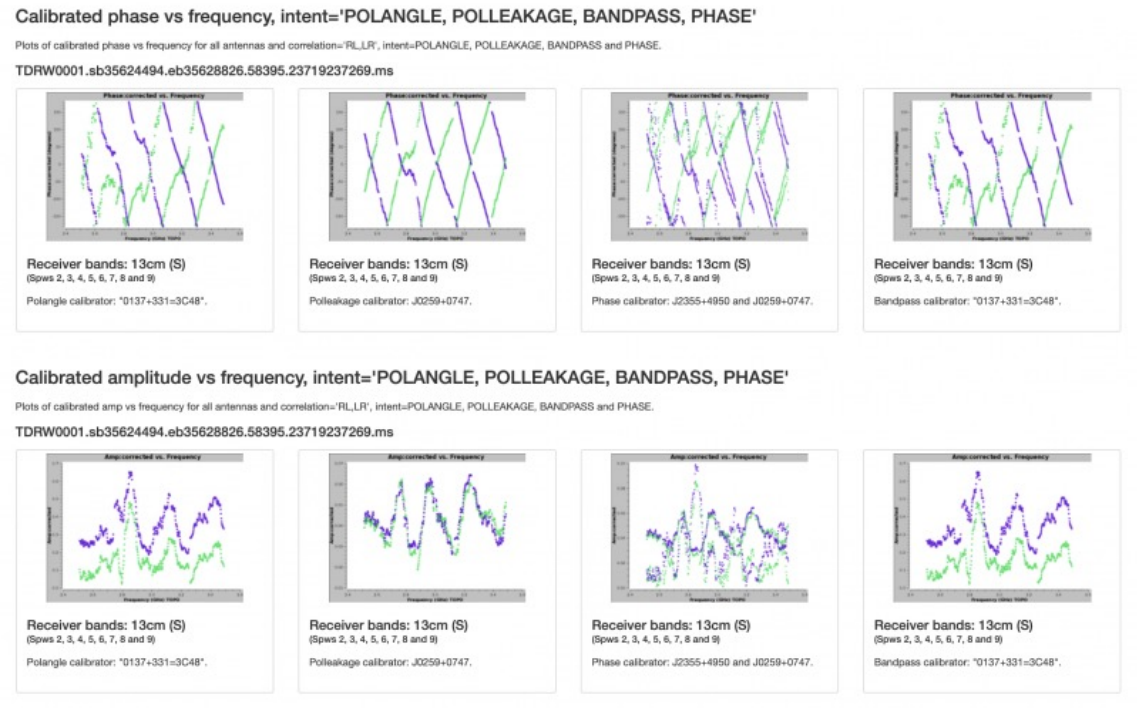
- A small number of pipeline parameters will be made available to tweak calibrations and images.
- For calibrations, spectral line avoidance (cont.dat) and RFI flagging, and optional calibration steps (TEC map and compression fix application) would be examples of things that could be changed.
- For imaging, the robust parameter/target resolution, frequency/velocity range and continuum ranges for spectra, and whether or not to attempt self-calibration.

Imaging from combined EBs

- The VLA lacks the hierarchical database structure that allows ALMA observations to be linked together. We will commission a user-driven imaging mode that allows different EBs from the same SB to be combined.
- This will later be extended to allow multi-configuration imaging using (user-specified) EBs from different configurations

Polarization calibration and imaging

- Polarization calibration will need to be applied by the user if that is required:
  - *https://casaguides.nrao.edu/index.php?title=Karl_G._Jansky_VLA_Tutorials#Polarization_Calibration_based_on_CASA_pipeline_standard_reduction:_The_radio_galaxy_3C75)*
- Pipeline will already produce plots of uncalibrated cross-hand phase and amplitude if polarization intents are detected.
  - Need SB info though to determine the intended polarization calibration strategy.



TTA tools

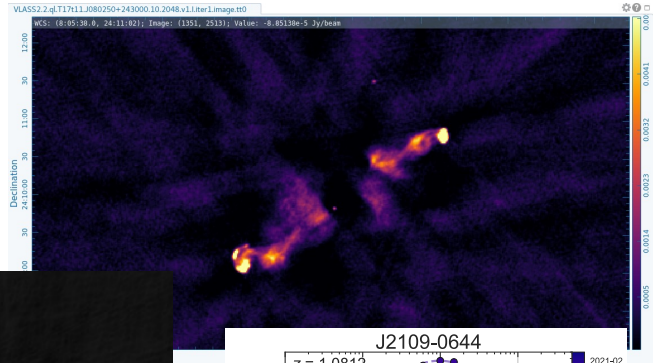
- A new suite of proposal, observing-block generation and scheduling tools is being developed.
- This will allow us to improve our SRDPs, for example:
  - Associate EBs that should be combined for imaging
  - Specify frequencies where strong spectral lines are likely to feed into the spectral line pipeline.
  - Specify the polarization calibration method (Df, Df+QU) so we can include polarization in SRDP.

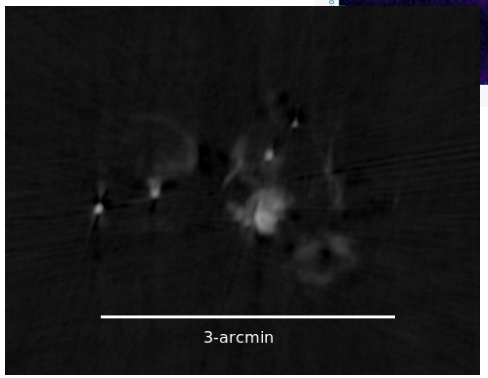
Pause for questions on future plans

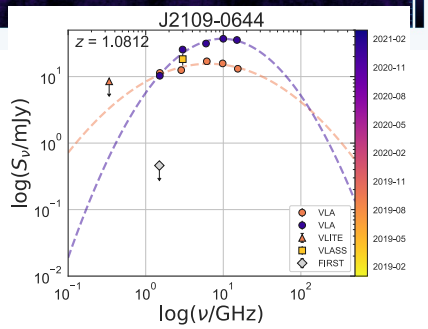
VLASS and VLASS data products

- What is VLASS? science.nrao.edu/vlass
  - All the sky north of dec -40 deg, imaged in three epochs ~32 months apart. "On-the-fly" raster mode observing.
  - S-band (2-4 GHz); 3" resolution; RMS~120-170μJy (RFI dependent).
  - IQU polarization (will also store some Stokes V information).
  - Radio astronomy community heavily involved in the survey from the design stage onwards.
  - All products are public following QA and archiving.
  - First epoch and second epoch now complete, observing for the third and final epoch begins in January 2023.
  - Brought under the SRDP program in 2019

VLASS and VLASS data products



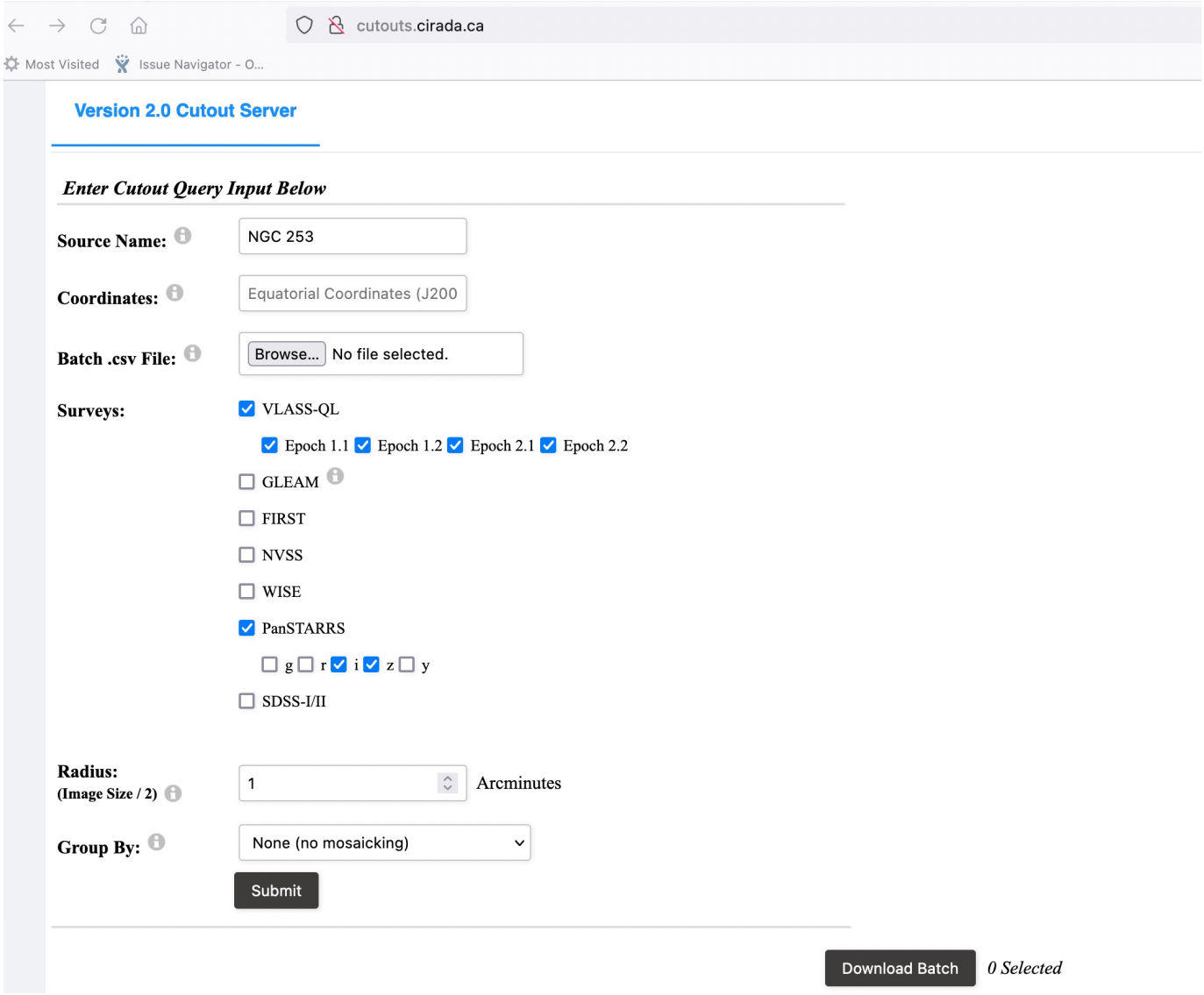




- Science:
  - Transients and variability
  - AGN and galaxy evolution
  - Polarimetry
  - Milky Way studies
- ``Quick Look" images are now available for both Epochs 1 and 2.
- More refined ``Single Epoch" images (including self-calibration, spectral index and polarization information) are starting to be produced now and will be made fully available over the next few years.
- At the end of the survey, a combined image will be made from all three epochs of data.

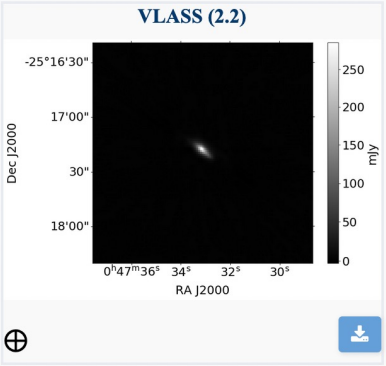
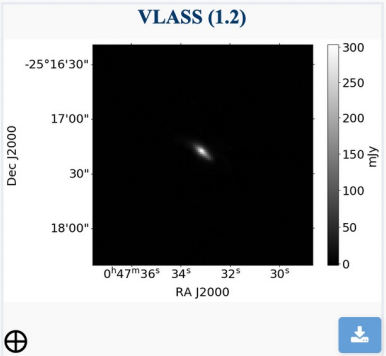
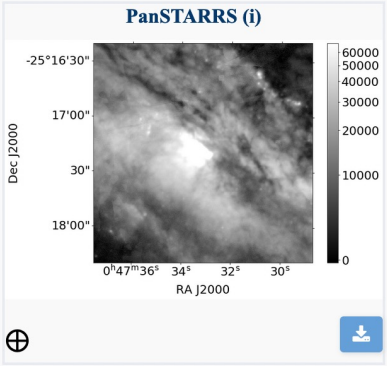
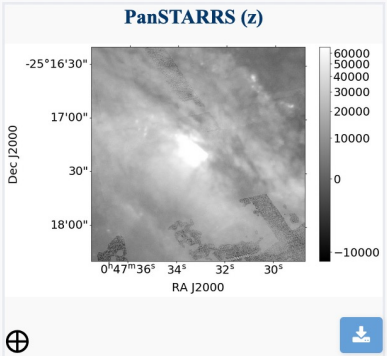
VLASS data products access

- VLASS raw data, calibrations and weblogs are available in the archive (data.nrao.edu) (images are available too, but not the final position corrected versions for Epoch 1).
- The latest VLASS images are available from archive-new.nrao.edu/vlass, and from CADC (https://www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca/en/vlass/). CADC services include VO SIA and TAP, and data can be accessed via astroquery.
- Our Canadian colleagues have also made a handy cutout server at http://cutouts.cirada.ca/
  - Makes cutouts of VLASS images, also NVSS/FIRST/GLEAM/RACS (soon) radio surveys, WISE infrared and SDSS/PanSTARSS optical.

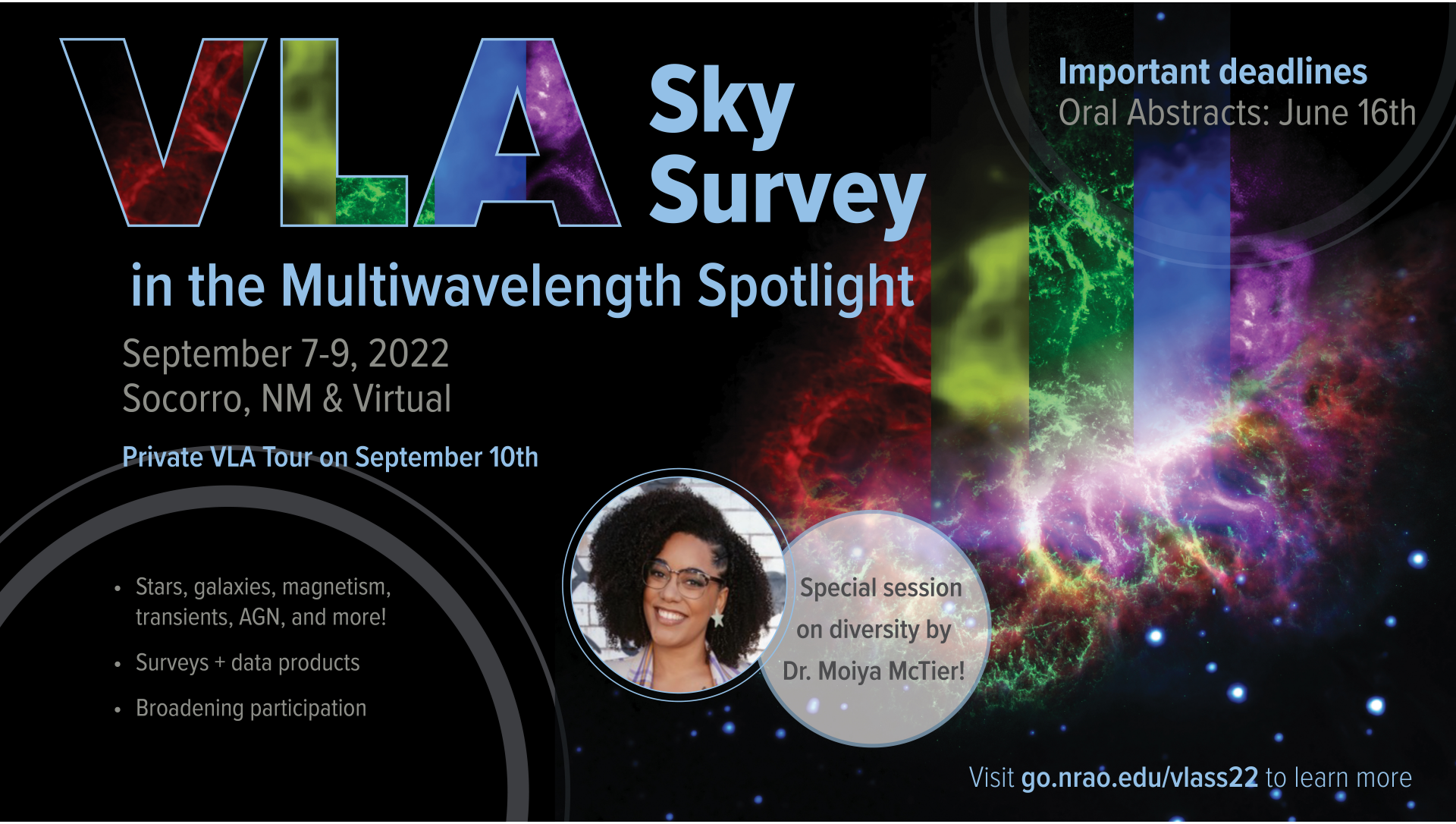

cutouts.cirada.ca
Most Visited
Issue Navigator - O...
Version 2.0 Cutout Server
Enter Cutout Query Input Below
Source Name:
NGC 253
Coordinates:
Equatorial Coordinates (J200
Batch .csv File:
Browse...
No file selected.
Surveys:
VLASS-QL
Epoch 1.1
Epoch 1.2
Epoch 2.1
Epoch 2.2
GLEAM
FIRST
NVSS
WISE
PanSTARRS
g
r
i
z
y
SDSS-I/II
Radius:
(Image Size / 2)
1
Arcminutes
Group By:
None (no mosaicking)
Submit
Download Batch
0 Selected

CIRADA cutout server

Showing 1 of 1 results

#	Input Name	RA (J2000) [deg]	Dec. (J2000) [deg]		VLASS	PanSTARRS
1	NGC 253	11.8880583	-25.2888000		 	 

VLASS22 conference



Lessons learned from SRDP

- Closely examining the data coming off the VLA has many benefits
  - Improved feedback to engineering/technical teams on issues with the telescope.
  - Detection of a flare in one of the ``big four'' calibrators, 3C138 in December 2020.
  - Identification of some issues with the calibrator models in CASA.
- In addition, the combination of VLASS and SRDP has pushed software development:
  - Fix for gain compression using the switched power data.
  - Improved workflows for calibration and imaging, ability to run these workflows on external clusters if necessary.

Summary

- We are already at the point where users are having trouble dealing with the volume of data from the VLA and ALMA, and the complicated algorithms needed to correctly process it.
  - With ngVLA and the ALMA correlator upgrade on the horizon, it will just get worse.
- SRDP delivers standard pipelined VLA data products that users can either use ``off the shelf'' or as a basis for more refined processing.
- From August, we will deliver images for VLA data that qualifies for SRDP.
- Still much to be done (spectral line processing, user-driven processing options, combined EB & combined array imaging etc).

www.nrao.edu
science.nrao.edu

The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.