Pipeline Calibration and Imaging for the Very Large Array

Brian R. Kent, Joseph S. Masters, Claire J. Chandler, John J. Tobin, Josh Marvil, Juergen Ott, Steven T. Myers, Drew Medlin, Amy E. Kimball, Frank K. Schinzel, Mark Lacy, Jeff Kern, Bryan J. Butler, Kana Sugimoto (NRAO), Dirk Muders (MPI Bonn), Stewart Williams, Vincent C. Geers (UK ATC)

Abstract

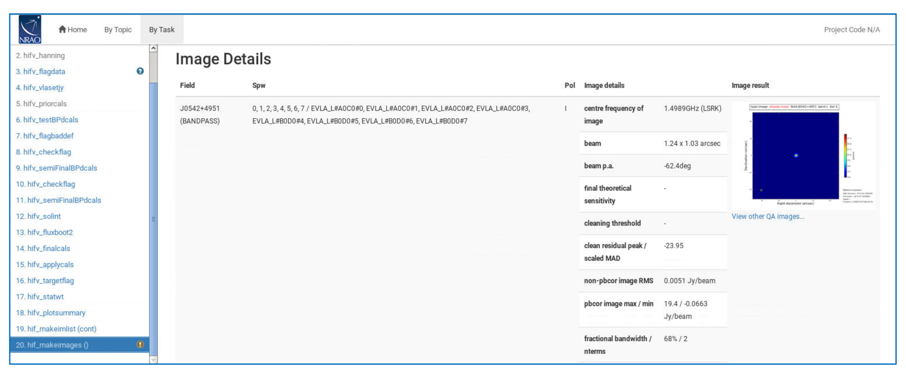
The VLA component of the CASA integrated pipeline produces calibrated measurement sets for the Karl G. Jansky Very Large Array. The pipeline is written in Python and consists of individual tasks and classes. Each task contains data processing heuristics designed to decide the best CASA task execution parameters for all VLA scheduling blocks above 1 GHz. The overall pipeline heuristics development has been improved through the Very Large Array Sky Survey (VLASS), a community and NRAO driven high-resolution 3 GHz continuum survey currently being carried out. VLASS produces imaging data products that are available to the astronomical community, and serves as a test platform for the NRAO's Science Ready Data Products initiative. We present the pipeline design and workflow as it pertains to the VLA and VLASS, calibration, imaging, and resulting products.

The pipeline is developed by an international consortium of scientists and software developers based at the National Radio Astronomical Observatory (NRAO), the European Southern Observatory (ESO), and the National Astronomical Observatory of Japan (NAOJ), and the Joint ALMA Observatory (JAO).

Design and Features

The pipeline consists of an execution framework written in Python. Its design is based around tasks, each with their own unit of functionality. Pipeline tasks are collected into serialized procedures and executed through a pipeline processing request XML (PPR-XML) or Python CASA script. As each stage is completed through the pipeline execution, a result object is generated with all associated information – task inputs, heuristic decisions, computations, and quality assurance (QA) information. These result objects are collected in the pipeline context, and updated as each subsequent stage is completed, maintaining the current state of the pipeline and allowing for breakpoints and restarts.

The pipeline weblog (below) acts as the user interface in understanding the results of the pipeline. Written as a responsive webpage with with Bootstrap and Python Mako templates, all relevant metadata concerning the observations, data reduction notifications, QA metrics, and heuristic outcomes are displayed for each task. The weblog is designed to assist in gauging the quality of calibration and imaging, so that a user may make the appropriate decisions with any additional data reduction requirements or additional reprocessing (for example, additional flagging or evaluating data for future observations).



See poster 10.41 by Masters et al. (2019) at this conference for descriptions of the ALMA pipeline.

Calibration and Imaging

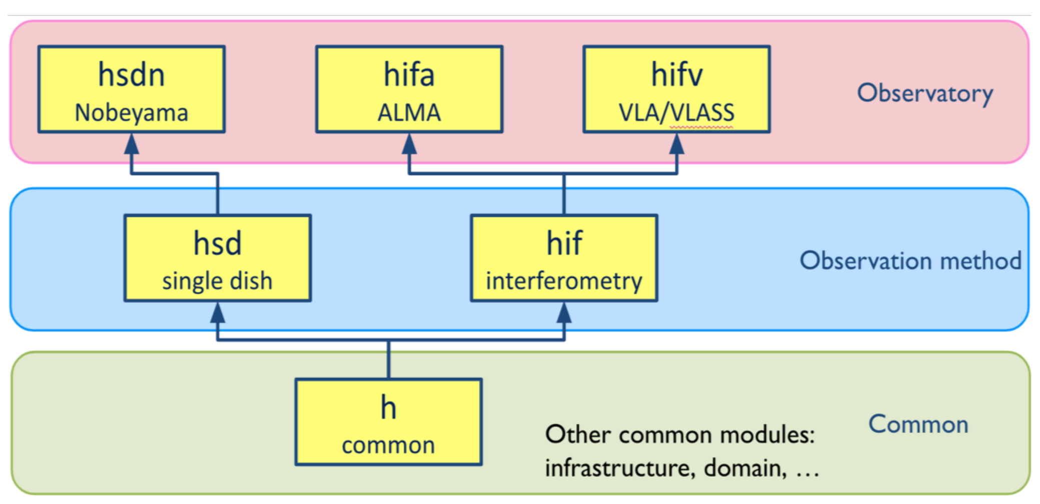
The standard VLA continuum data reduction procedure determines, on the basis of a priori factors and from observations of standard calibration sources, the corrections to the raw data amplitude, phase, and visibility weights to be applied to the data. This process also determines the flags that are needed to remove bad data due to instrumental faults, RFI, and other causes of error. This process only includes the derivation of the complex gain and bandpass calibration factors known through previous measurements or determined by the observations of calibrators and transferred to the target observations.

Calibration stages consist of:

1. Import from SDM to measurement set and application of initial online flags (off-source, focus error, subreflector error)
2. Determination and application of derived flags (RFI, bad antennas, shadowing, other)
3. Switched power amplitude calibration and antenna gain curves
4. Flux scale calibration (using standard sources)
5. Complex Delay and Bandpass Calibration, Complex Gain Calibration, followed by additional heuristic flagging
6. Flux density bootstrapping (from primary to secondary calibrators)
7. Interpolation and Application of Cumulative Calibration
8. Final Flagging of Data (insufficient or failed calibration, RFI) and statistical weighting of visibilities
9. Diagnostic calibrator imaging
10. Output of Quality Assurance (QA) information, plots, and images

In addition, the standard VLASS calibration procedure included different steps for heuristic flagging with RFLAG and TFCROP, as well as polarization calibration. The diagram below shows the inheritance design for the pipeline. See the VLA science website for information on executing the pipeline:

https://science.nrao.edu/facilities/vla/data-processing/pipeline/



References

Condon, J. J. et al. 1998, AJ, 115, 1693
Lacy, M. et al. 2019, VLASS overview (arxiv: 1907.01981)

The VLA Sky Survey

The VLA Sky Survey (VLASS, Lacy et al. 2019), a seven year community-driven endeavor started in September 2017 is currently surveying the entire sky down to a declination of -40 degrees at S-band (2-4 GHz). This 5500 hour next-generation large radio survey will explore the time and spectral domains, relying on pipeline processing to generate calibrated measurement sets, polarimetry, and imaging data products that are available to the astronomical community with no proprietary period. The project will complete the all-sky survey in three epochs. The source extraction effort is expected to catalog approximately 10 million radio sources. The high resolution images will uncover old and new or short-lived (transient) objects and show distant radio galaxies in unprecedented detail.

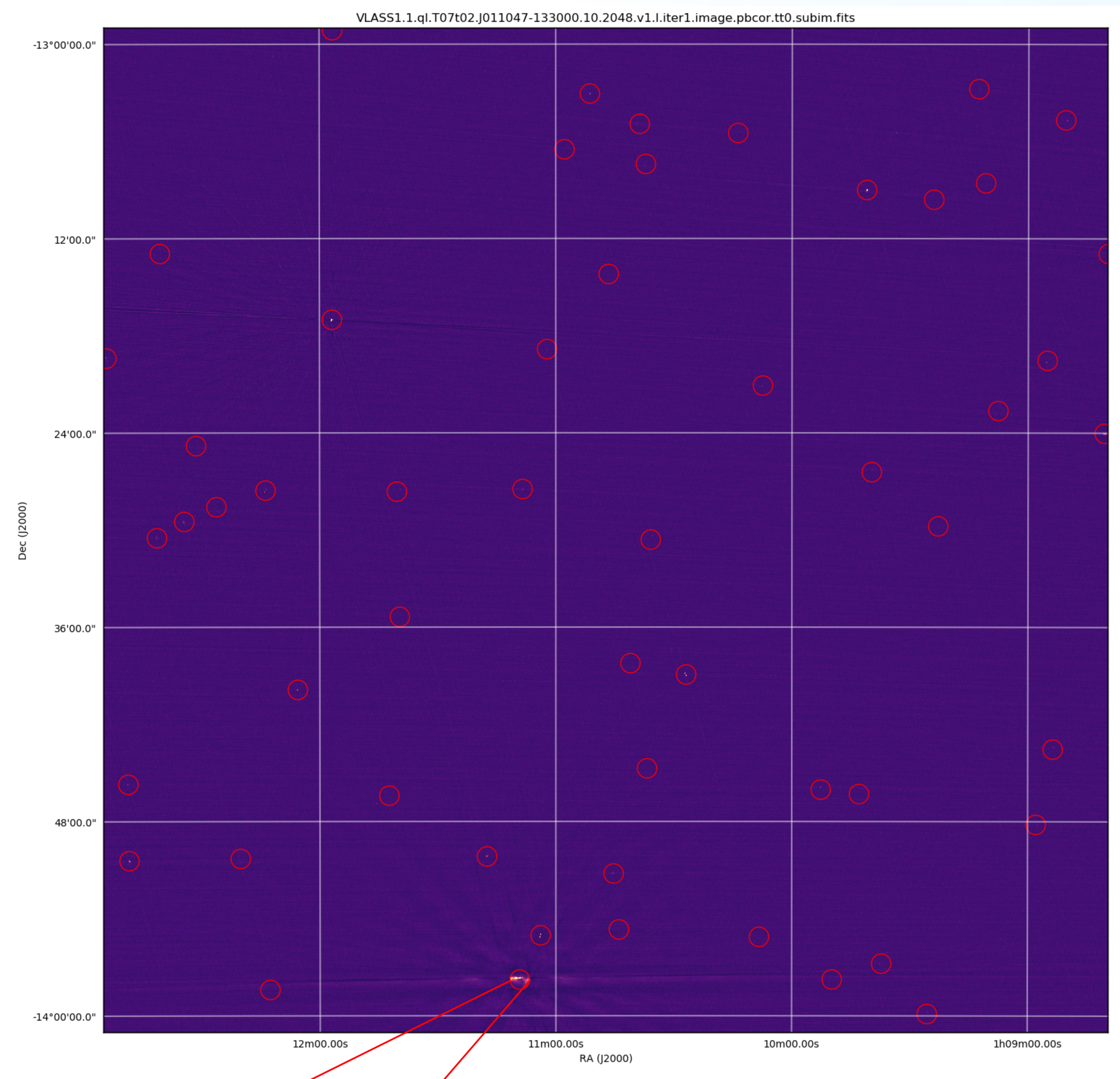
https://public.nrao.edu/vlass/

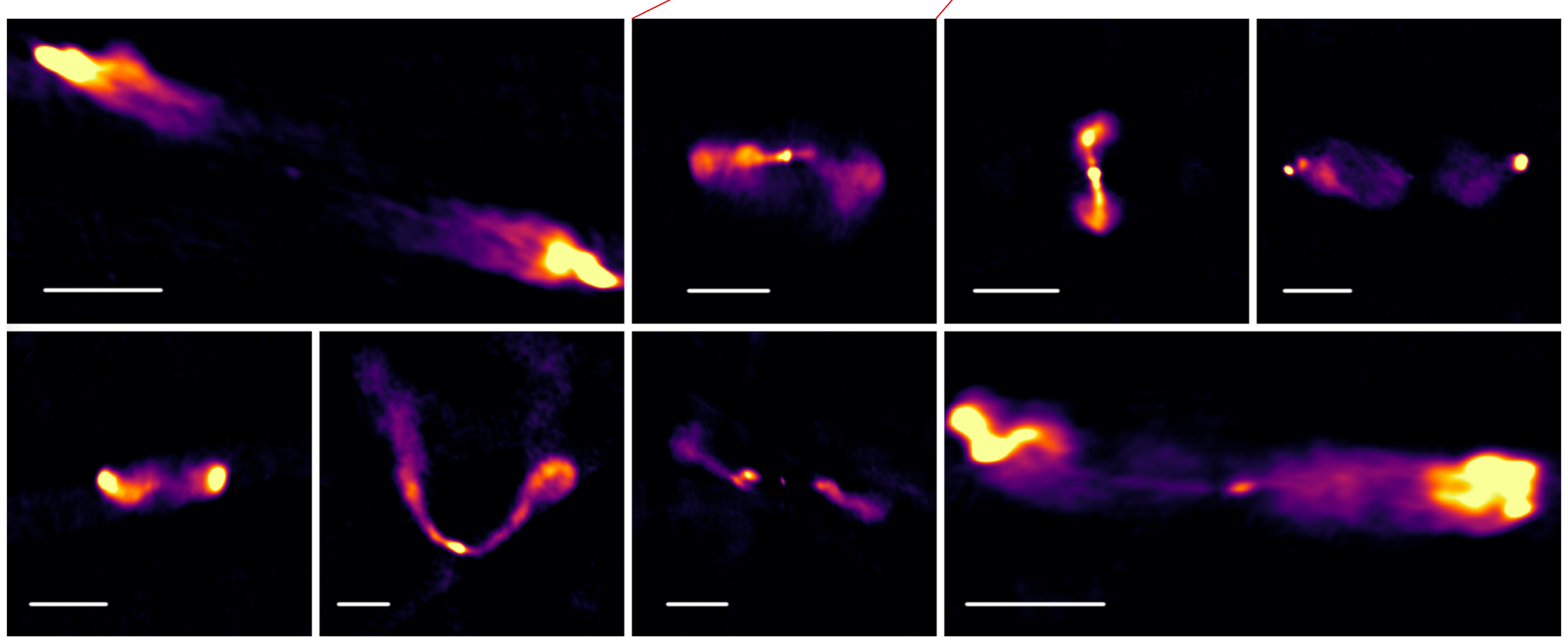
The image at right shows a 1 x 1 degree quicklook cutout image. The red circles identify previous detections with the L-band NVSS survey (Condon et al. 1998).

VLASS data products are available via:

https://archive-new.nrao.edu/vlass/

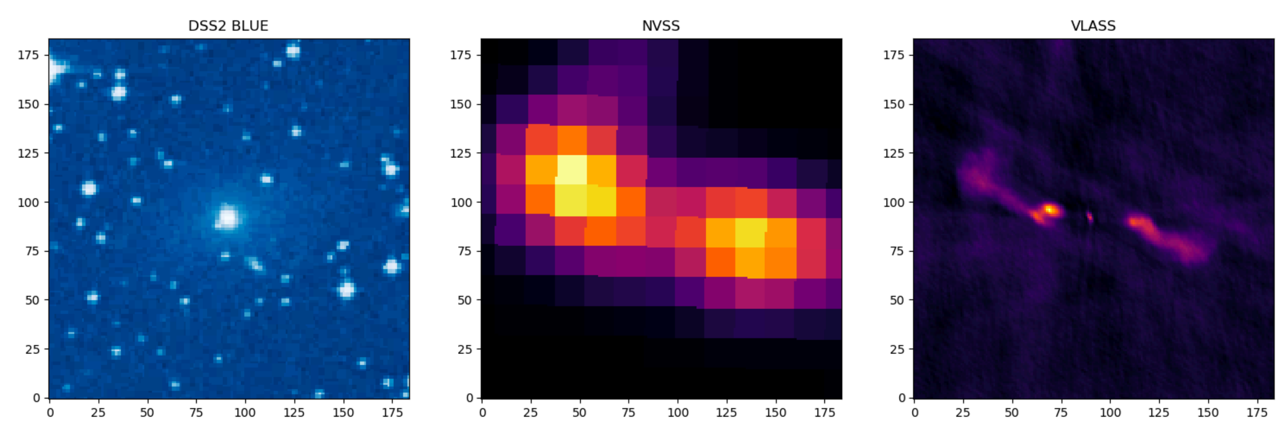
The Very Large Array Sky Survey	
Frequency	2 – 4 GHz
Resolution	2.5 arcsec
Sky coverage	All sky north of Dec > -40 deg (33885 sq. deg.)
Sensitivity	120 microJy RMS
Combined (3 epoch sensitivity)	69 mJy RMS
Polarization	I, Q, U
Cadence	3 epochs separated by 32 months
Start Date	September 15, 2017
Expected source count	~ 10 million





Cutouts from VLASS quicklook image data products. Each white scale bar is 30 arcseconds. The images depict previously unresolved radio galaxies of different morphologies.

Comparisons made of an optical image from the Digital Sky Survey, original L-band D-configuration NRAO VLA Sky Survey (NVSS), and a three arcminute quick look cutout from VLASS. Both axes indicate arcseconds.



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