

NRAO Science Ready Data Products

Mark Lacy



What is the SRDP Project?

- Data from modern radio interferometers such as the VLA and ALMA is both very large in terms of volume, and complicated in terms of what the data model allows.
- Significant barriers now exist for newcomers to data from these instruments, and even individuals with expertise find data processing very burdensome.
- This reduces the scientific output of these facilities – rather than being limited by the technical capabilities of the instruments, scientists are limited by the logistics of data processing

SRDP

Aims to deliver data products that have been produced by an observatory-standard pipeline, quality-assured to a consistent standard, and have bad data removed.

- Overall plan:
 - Five Waves (waterfall development) expanding from basic capabilities to a full program over five years.
- Initial pilot & Wave 1:
 - Apply calibration tables to previously calibrated VLA and ALMA raw data (working).
 - Produce calibrated visibility (uv) data for VLA high frequency observations (working).
 - Produce bespoke images from calibrated ALMA visibilities to include only the channel range needed by the user (delayed).
 - Ingest products from large programs such as the VLA Sky Survey

Waves 2-5

- Future waves will include imaging of VLA data, special workflows for time-critical VLA observations, and multi-configuration/array imaging for VLA and ALMA.
- Reruns of pipeline calibrations with user-tuned parameters for VLA and ALMA will be allowed.
- We also will include the capability to ingest data products from large VLA and ALMA programs into the archive and serve them to the community.
- A new system for observing proposals and the generation of scheduling blocks will ensure that observations are more compatible with the pipeline's expectations regarding calibration strategies etc.

Resources

- SRDP has to be carried out so far as possible within the Observatory's existing resources of compute and disk space.
- The limited scope of the early waves will mean this is not a problem.
- Later waves, that will include VLA imaging, may be more problematic and require additional resources.
- Working to make the CASA pipeline more compatible with Open Science Grid and other computing resources beyond NRAO's clusters.

The SRDP pilot

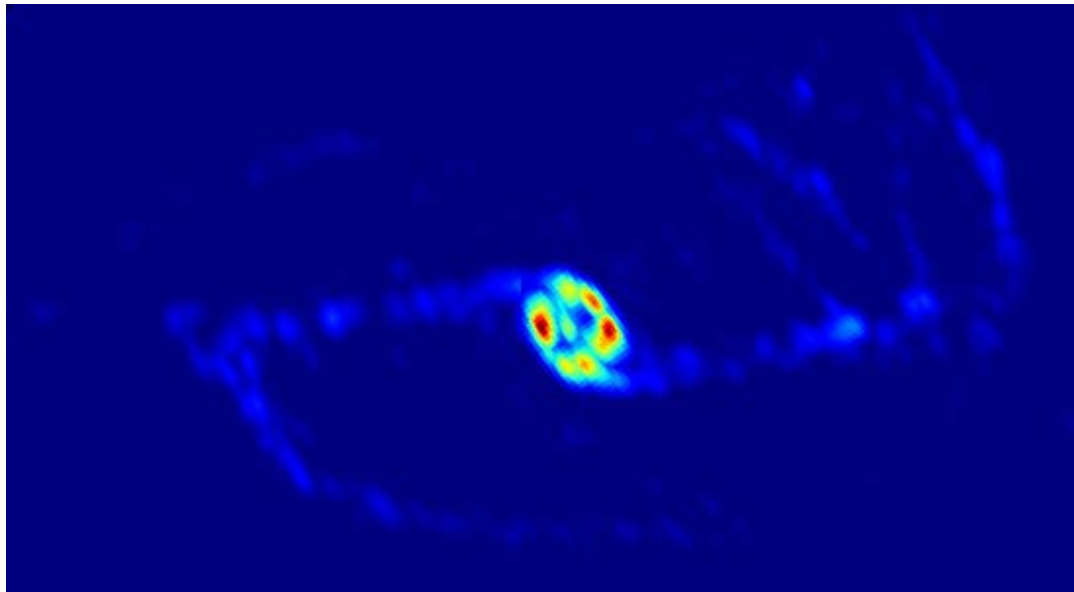
- VLA calibration began in mid-June 2019
 - Initially restricted to high frequencies (where interference is negligible) $> 12\text{GHz}$ (Ku-band and above), and to single-band datasets that followed standard calibration procedures.
 - Included X-band (8-12 GHz) in September.
- Procedure:
 - Standard pipeline run on data (as usual)
 - Results reviewed by a data analyst, and suggestions for flagging or other changes made.
 - Suggestions are reviewed by an “Astronomer on Duty” (AoD) (a scientist or an experienced analyst).
 - Pipeline is rerun.
 - Process above is iterated until AoD signs off on a QA pass or fail.

VLA calibration

- Since the start of the pilot on June 15th, we have calibrated 61/63 qualifying datasets to SRDP standards (about 10% of the total observed).
- The average number of times a datasets needs to be rerun through the pipeline is 1.1.
- The median time to delivery from observation (including the initial pipeline run) is 5 days.
- These results show that the SRDP calibration project will be feasible, and the limited feedback we have had so far suggests our users are happy with the products.
- Pipeline improvements scheduled for Wave 2 (RFI flagging, gain table flagging and ability to deal with multi-band datasets) will allow us to expand to more of the data (including lower frequencies), and reduce the number of reruns required.

ALMA Imaging (currently under test)

- Users can select their own imaging parameters via a web interface
- Pipeline software will apply the calibration to the raw data, then make the image per the users' request.
- Allows the user to image the part of the cube they want, at the resolution they need.



Launch Optimized Imaging on: 2017.1.00236.S



User Email (required):

Request Description:

Request Description

Destination Directory:

Specify directory (must be logged in & staff)

/lustre/

CASA Version:

5.4.2-8 (recommended) ▾

SPW:

spw:19 (225.392 GHz-227.330 GHz) dnu = 124995k...

Field:

ngc_2992

Frequency Space

Velocity Space

Rest Frequency:

GHz

Start:

km/s

Width:

km/s

N Channels:

8

End:

km/s

Cancel

Submit Request

Title: Feeding and feedback in an unbiased an
Abstract: We will map CO(2-1) in the central 3
 quantify the global AGN feeding and feedback
 down to 10^5 Mo, derive outflow rates, gas re
 measure inflow rates. With the continuum map
 ALMA archival observations, will provide a san
 legacy value comes from combining uniform C
 resolution follow-up to dynamically constrain t
PI: Matthew Malkan

Co-Authors: Francoise Combes, Miguel Perei
 Masatoshi Imanishi, Matthew Malkan, Kotaro h

MOUSes

Images

MOUS

ngc_2992_a_06_TM1



om the 127m all-sky sample. This is crucial to
 xy evolution. We will detect molecular outflows
 he gravitational torques driving infall, and
 2 bright AGN in 27 hours which, when added to
 of a similarly selected sample of LINERs. The
 itable bright nearby AGN for ALMA high-

lio Ricci, Jason Glenn, Paola Michela Andreani,

ownload Restored MS

Optimized Imaging

National Science Foundation operated
 d Universities, Inc.



science.nrao.edu
public.nrao.edu
ngvla.nrao.edu

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