

VLA Pipeline(s) & SRDP

Anna D. Kapinska (NRAO)
Slides courtesy of Drew Medlin (NRAO)







VLA Calibration Pipeline – Overview

- With the start of (Jansky) VLA Full Operations (January 2013), pipeline automatically run on all Scheduling Blocks as soon as the data are ingested into the archive (over 24,700 to date):
 - Deliver flagged and calibrated visibility data
 - You will self-calibrate and image visibility data to meet science goals, using resources at home institution or NRAO computing resources
- Automated pipeline should run correctly on all "standard"
 Stokes I science SBs; "standard" means:
 - 128 MHz spws, but may work on other set-ups as well
 - Some constraints on strength of calibrators needed
 - Contains correctly labeled and complete scan intents
 - → at the minimum: flux density calibrator, complex gain calibrator



VLA Calibration Pipeline – Overview

The pipeline successfully completes on ~96% of all science SBs observed on the VLA; whether the output can be used for science depends on the science goal, and whether the observations were correctly set up

Caveats:

spectral line projects, & polarisation observations

- Current versions available:
 - CASA integrated pipeline: compatible with ALMA pipeline infrastructure, and used as real-time pipeline since Sep 2015
 - "scripted" pipeline: collection of python scripts that use CASA tasks wherever possible, but also uses toolkit calls; readable and easy to modify. It was the original VLA pipeline and in use in realtime pipeline operations from early 2013 and until Sep 2015.

VLA Calibration Pipeline – Overview

- Real-time pipeline at NRAO:
 - Minimal human intervention: Pipeline is run automatically on every science SB as it completes (not just "continuum")
 - Pipeline output undergoes basic quality assurance checks by NRAO staff, and detailed checks are made for most C-band and higher continuum; reports generated are archived as pipeline products
- At your home institution:
 - Instructions for installation and operation of the VLA CASA Calibration
 Pipeline are available at http://go.nrao.edu/vla-pipe
 - Uses CASA 6.2.1, similar to current real-time pipeline
 - See the VLA CASA pipeline guide at http://go.nrao.edu/vla-casa-tut



Pipeline Heuristics

- Assuming requirements are met, the pipeline:
 - Loads the data (SDM-BDF \rightarrow MS)
 - Hanning smoothing*
 - Retrieves information about the observing set-up from the data
 - Applies deterministic flags (online flags, shadowed data, end channels of spectral windows, etc.)
 - Identifies primary calibrators and loads models

^{*}May want to modify inputs and/or omit entirely for spectral line reductions, unless heavily impacted by RFI or dealing with a very strong spectral line feature.

Pipeline Heuristics

- Derives all prior calibrations (antenna position corrections, gain curves, atmospheric opacity, requantizer gains)
- Iteratively determines initial delay and bandpass solutions, including running RFLAG, and identifying system problems
- Derives initial gain solutions, does flux density bootstrapping and derives spectral index of all calibrators, sets models.
- Derives final delay, bandpass, and complex gain calibrations
- Applies all calibrations to the MS
- Runs RFLAG algorithm on all fields, including target*
- Runs statwt to derive proper relative weights per antenna/spw*







♠ Home

y Topic

Project Code N/A

Observation Overview

Project	uid://evla/pdb/35621723
Principal Investigator	Dr. Emmanuel Momjian
Observation Start	2018-10-04 05:41:35 UTC
Observation End	2018-10-04 08:32:45 UTC

Pipeline Summary

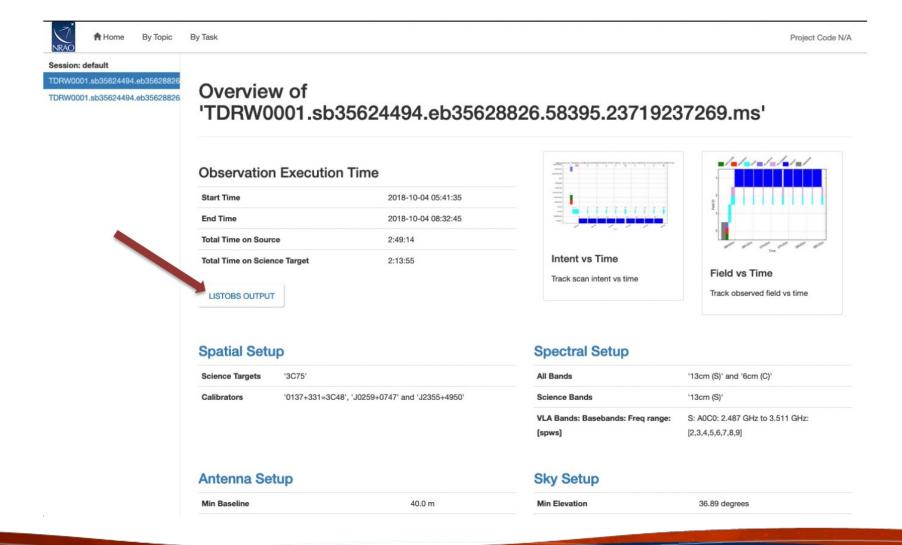
Pipeline Version	2021.2.0.128
CASA Version	6.2.1.7 (environment)
IERSeop2000 Version	0001.0151 (last date: 2021-08-01 00:00:00)
IERSpredict Version	0623.0600 (last date: 2022-01-01 00:00:00)
Pipeline Start	2021-11-16 15:08:01 UTC
Execution Duration	4:31:14

Observation Summary

			Time (UTC)			Baseline Length			
Measurement Set	Receivers	Num Antennas	Start	End	On Target	Min	Max	RMS	Size
Scheduling Block ID: uid://evla/pdbsb/35624494									
Session: default									
TDRW0001.sb35624494.eb35628826.58395.23719237269.ms	13cm (S)	27	2018-10-04 05:41:35	2018-10-04 08:32:45	2:13:55	40.0 m	1.0 km	441.9 m	11.7 GE
TDRW0001.sb35624494.eb35_18826.58395.23719237269_target.ms	13cm (S)	27	2018-10-04 06:04:00	2018-10-04 08:30:00	2:13:55	40.0 m	1.0 km	441.9 m	9.2 GB

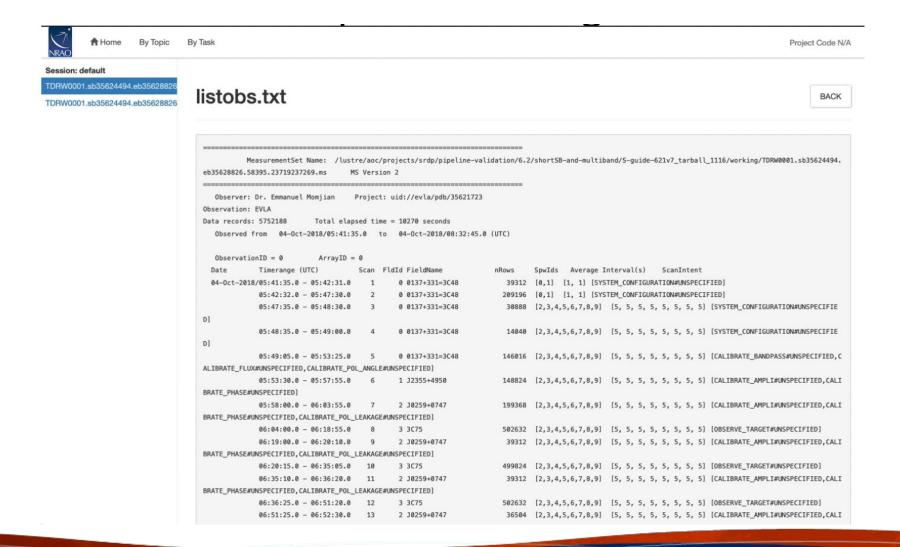


















Task Summaries

Took	QA Score		Duration
1. hifv_importdata: Register VLA measurement sets with the pipeline		1.00	0:08:34
2. htfv_hanning: VLA Hanning Smoothing	A	N/A	0:03:42
3. hifv_flagdata: VLA Deterministic flagging		1.00	0:06:18
4. htfv_vlasetjy: Set calibrator model visibilities		(1.00)	0:01:35
5. htfv_priorcals: Priorcals (gaincurves, opacities, antenna positions corrections, rq gains, and switched power)	A .	N/A	0:27:51
6. hifv_testBPdcals: initial test calibrations		1.00	0:05:35
7. hifv_checkflag: Checkflag summary		0.97	0:02:02
8. htfv_semiFinalBPdcals: Semi-final delay and bandpass calibrations		1.00	0:05:23
9. hifv_checkflag: Checkflag summary		1.00	0:06:41
10. Nifv_solint: Determine solint and Test gain calibrations		1.00	0:03:01
11. htfv_fluxboot: Gain table for flux density bootstrapping		0.96	0:03:40
12. Nt/v_finalcals: Final Calibration Tables		1.00	0:09:42
13. htfv_applycals: Apply calibrations from context		1.00	0:05:20
14. htfv_checkflag: Checkflag summary		1.00	0:36:18
15. Nfv_targetflag: Targetflag		1.00	0:06:40
16. htfv_statwt: Reweight visibilities		1.00	0:13:01
17. htfv_plotsummary: VLA Plot Summary		1.00	0:09:17
18. Nf_makeimilist: Set-up parameters for phase calibrator & bandpass calibrator imaging	9	1.00	0:00:35
19. Nif_makeimages: Make calibrator images		1.00	0:08:05
20. http_exportdata: Prepare pipeline data products for export		1.00	0:02:27
21. Mf_mstransform: Create science target MS		1.00	0:05:30
22. hif_checkproductsize: Check product size		1.00	0:01:45
23. Nif_makeimlist: Set-up parameters for target aggregate continuum imaging	Ų.	1.00	0:01:26
24. htf_makeimages: Make target aggregate continuum images		1.00	1:33:50
25. htfv_pbcor: Pbcor	NA .	(N/A)	0:00:04
26. http_exportdata: Prepare pipeline data products for export		1.00	0:02:40





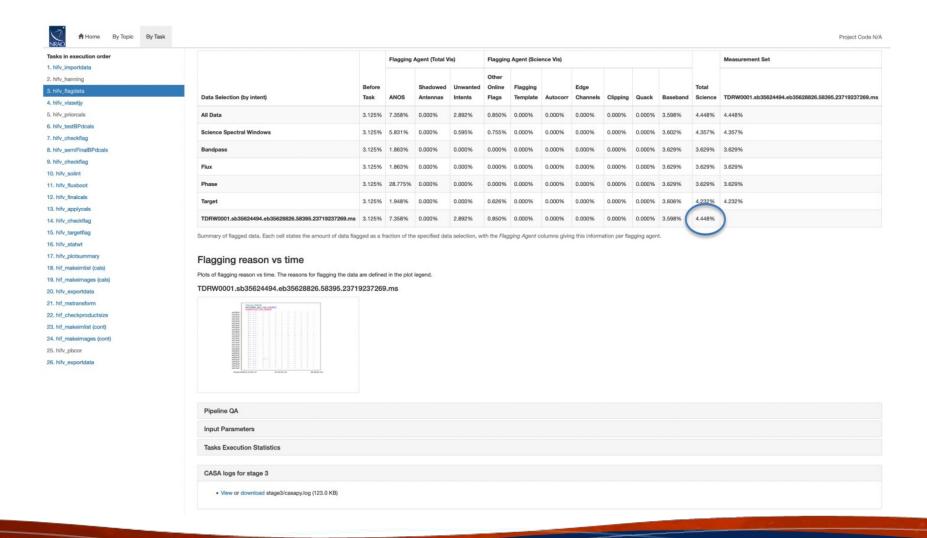
Project Code N/A

The following pipeline steps provide key checks for calibration quality:

```
hifv flagdata
                    deterministic flagged data fraction
hifv_testBPdcals
                    hardware problems and other obs. issues
hifv_solint
                    solution intervals for phase cals, input gain tables
hifv_fluxboot
                    fitted calibrator flux densities and spectral indices
hifv_finalcals
                    final calibration tables applied to the data
hifv_plotsummary
                     useful diagnostic plots of calibrated data
```

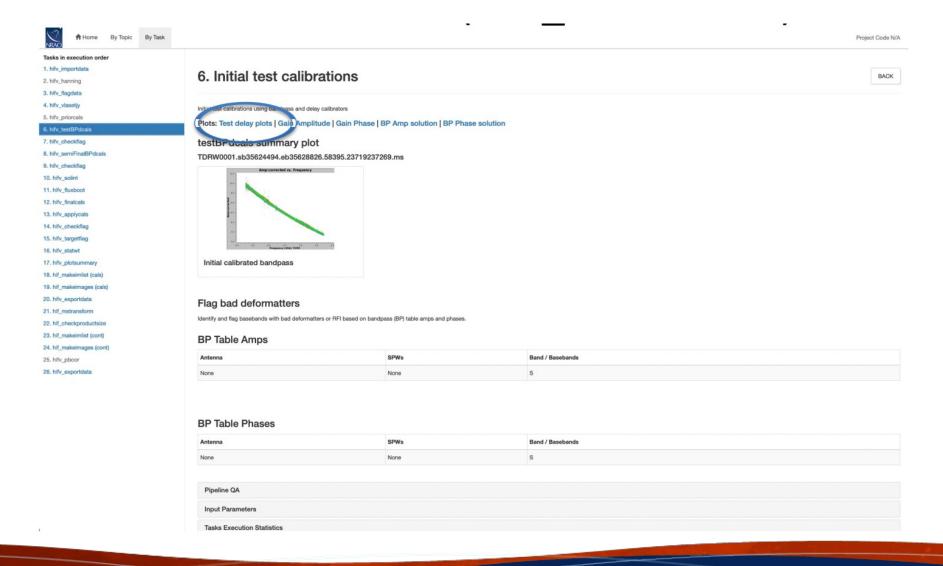


Deterministic Flagging (hifv_flagdata)



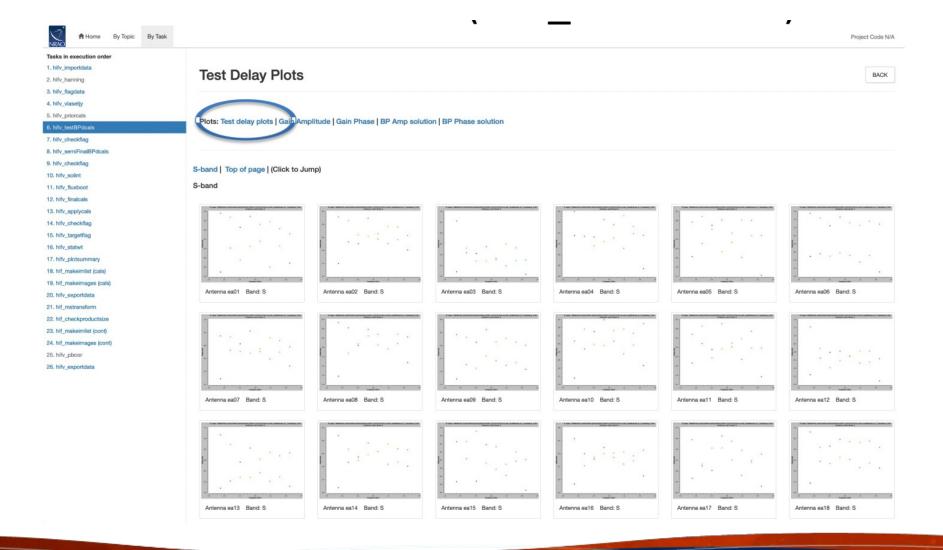






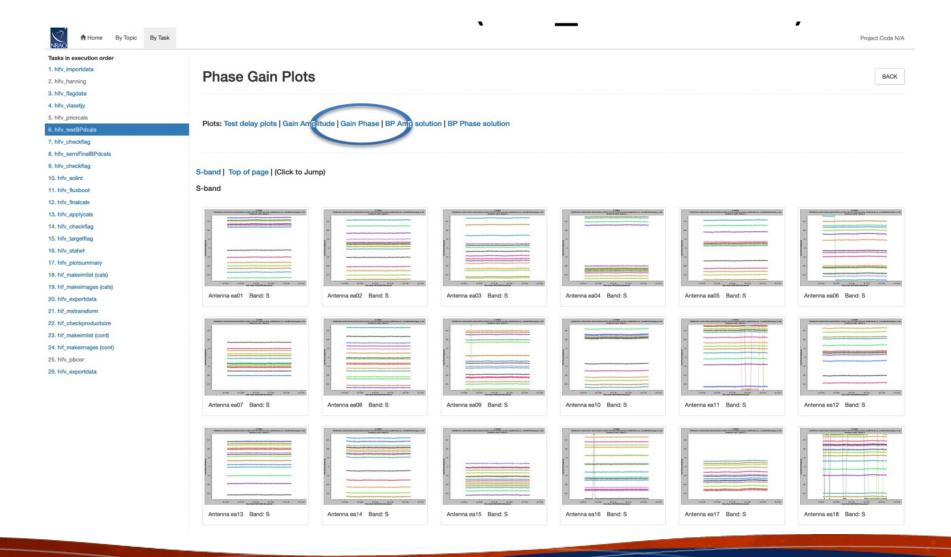






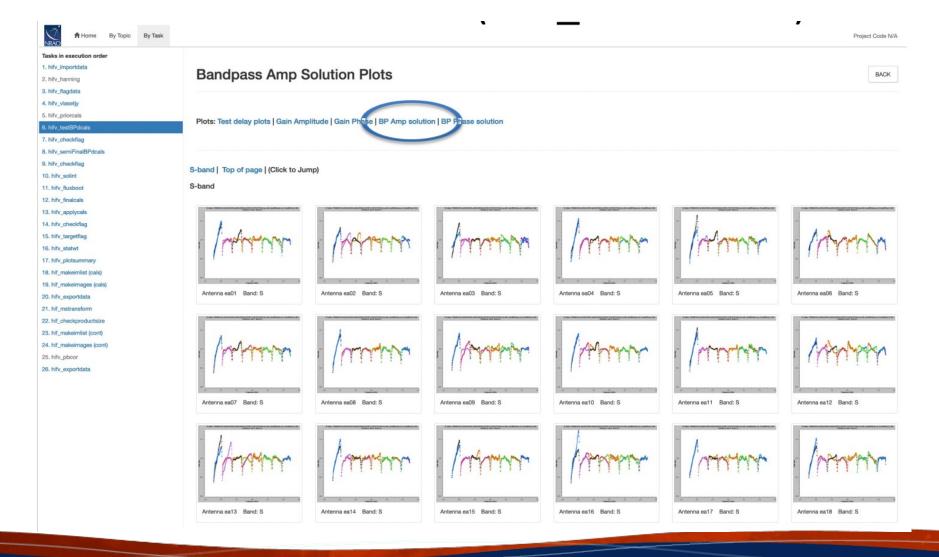








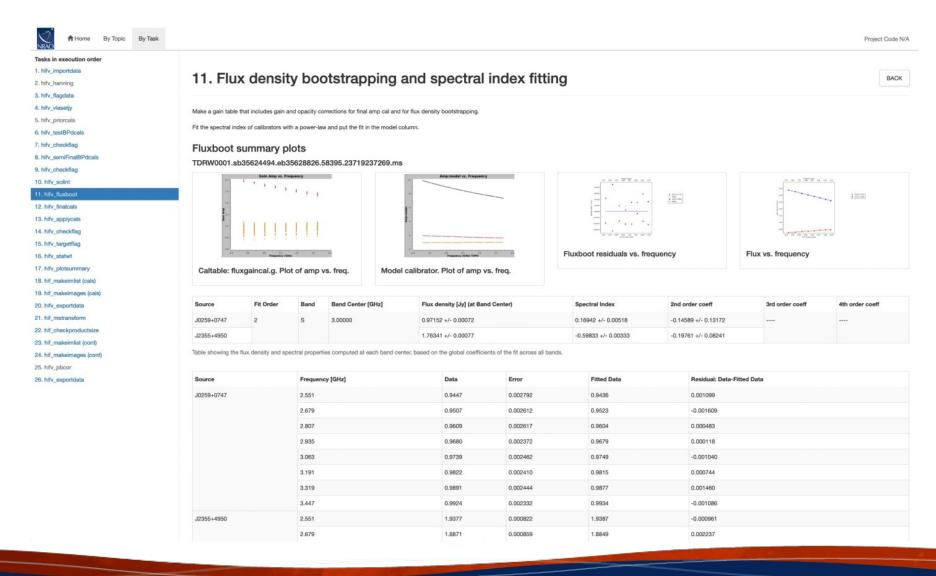








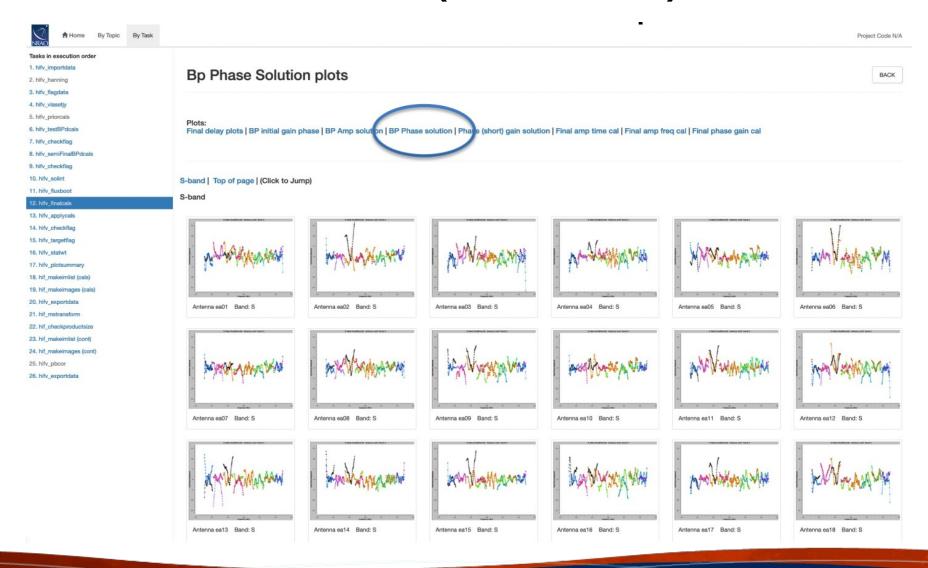
Flux density bootstrapping (hifv_fluxboot)







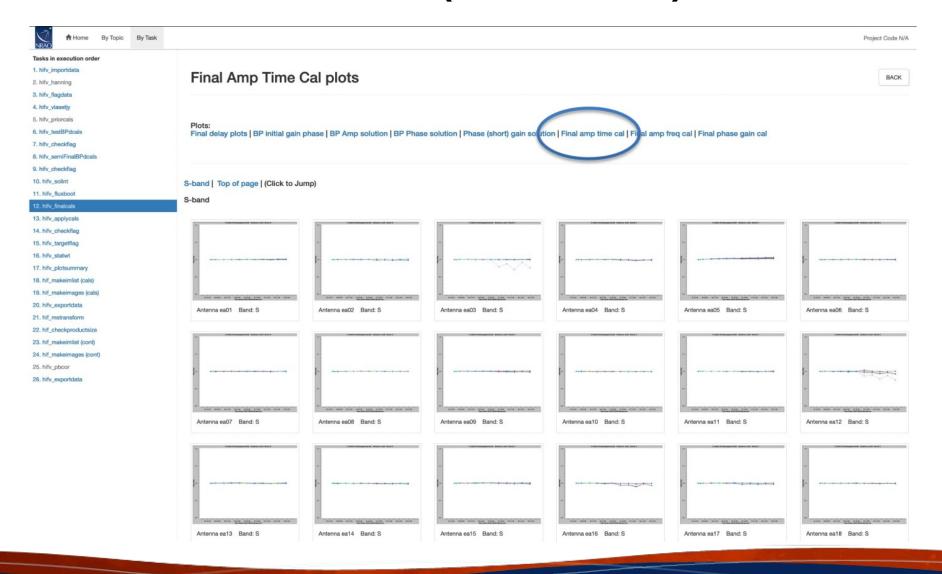
Final calibration tables (hifv_finalcals)







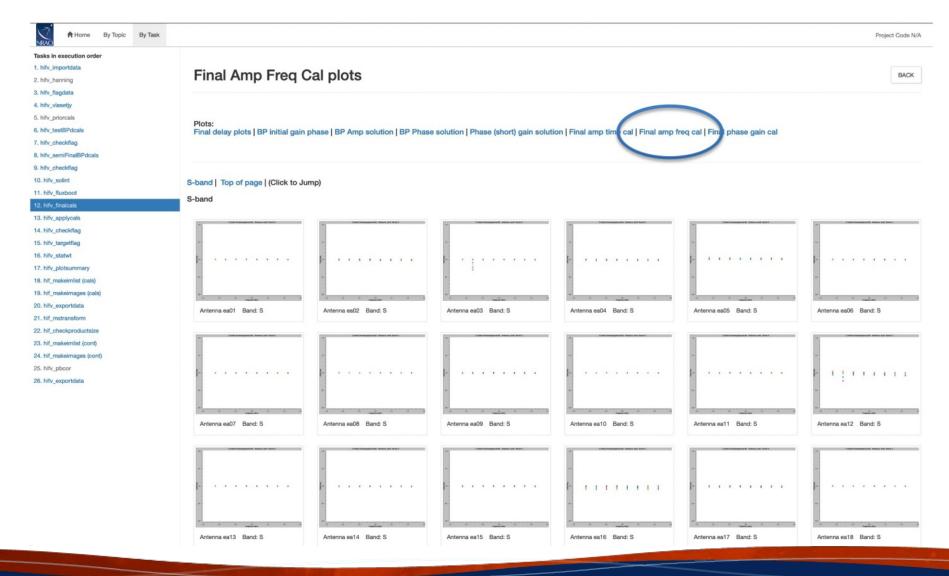
Final calibration tables (hifv_finalcals)







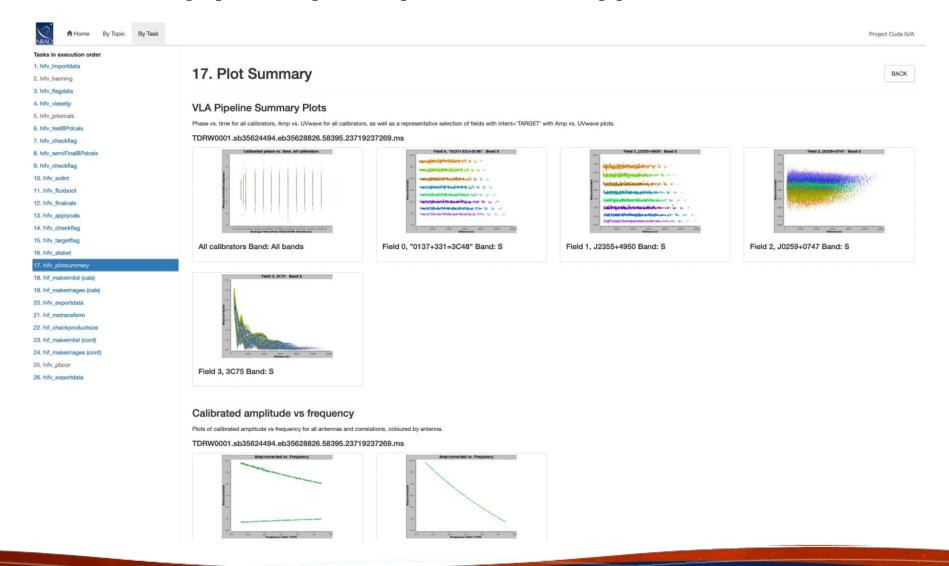
Final calibration tables (hifv_finalcals)







Summary plots (hifv_plotsummary)







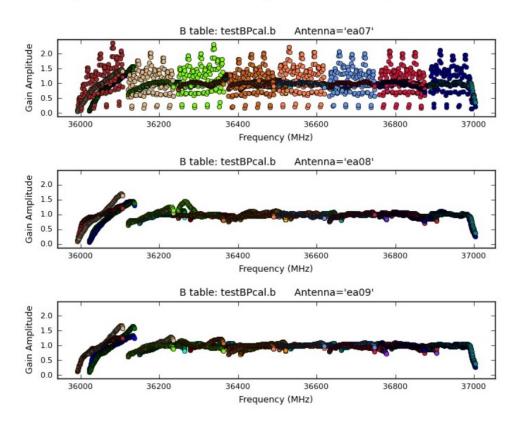
Known failure modes

- In general the pipeline does very well, but there are possible failure modes:
 - No flux density or gain calibrator intents defined, or flux density calibrator not one for which we have models
 - work around in scripted pipeline
 - Wrong scan intents
 - modify Scan.xml in SDM; see https://science.nrao.edu/facilities/vla/dataprocessing/pipeline#section-28
 - Does not always identify deformatter problems (but does NOT usually have false positives – L-band an exception)
 - flag remaining bad spws
 - Calibrators are too weak for given spw bandwidth
 - heuristics have been developed and are being tested



Examples of some issues

Digital transmission system ea07 bandpass, bad data (DTS issue); ea08, ea09 OK

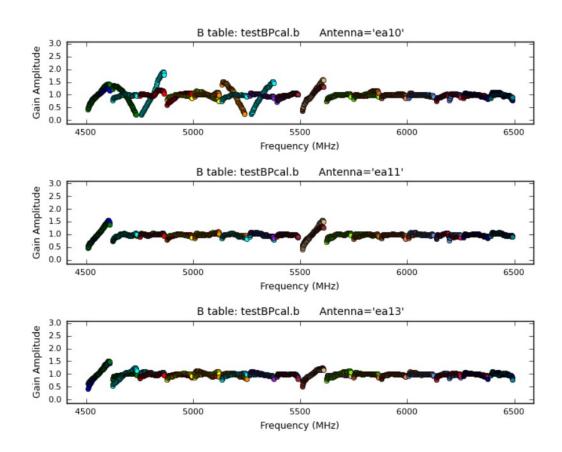






Examples of some issues

ea10 bandpass, bad data (DTS issue); ea11, ea12 OK

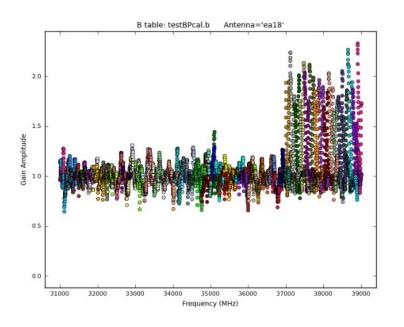


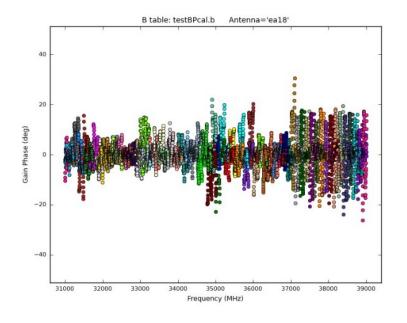




Examples of some issues

ea18 Amp and Phase affected (DTS issue for 37-39GHz)









Pipeline products and outputs

(1) Flagged and Calibrated MS (2) Final flag version and calibration tables (archived) (3) Logs, incl. weblog used by quality assurance (QA) staff and QA report (archived)

The real-time pipeline produces a calibrated and flagged MS

- → Calibrated MS may be requested through the archive https://data.nrao.edu
- → You may request a more detailed QA2 report from the Data Analysts (https://help.nrao.edu/, Pipeline Department)
- → If you are happy with the pipeline calibration, then:
 - Do further flagging if necessary
 - Split out your target and image (imaging pipeline now available)

If you have the SDM or uncalibrated MS and the calibration and flag tables, then the pipeline calibration can be restored, see details at

http://go.nrao.edu/vla-pipe/

Science Ready Data Products (SRDP)

- → calibrated & imaged science ready data
- → initiative started in June 2019

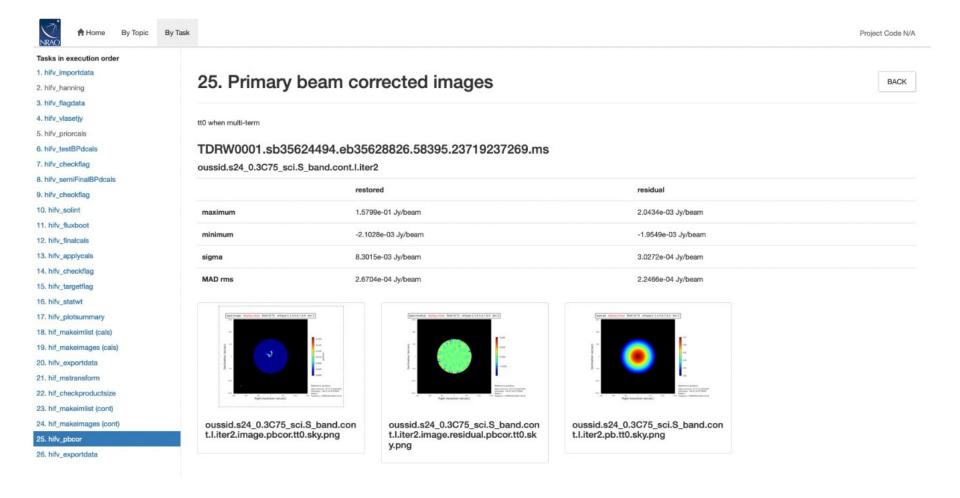
https://science.nrao.edu/srdp/home

In early stages, at the moment supporting only:

- Continuum (Stokes I)
- C-band or higher frequency
- if SB setup and scan intents set correctly for pipeline
- only observations using 3C286 or 3C147 as flux density calibrators

Quality Assurance: NRAO Staff will check quality in detail, add extra flagging and rerun if necessary.

SRDP: VLA Imaging Pipeline







Final notes

Viewing weblogs

The default security preferences in Firefox block weblogs on disk from being viewed directly

→ Need to change about:config: security.fileuri.strict_origin_policy to False

Starting the pipeline

Start with SDM-BDF raw data (MS possible, but online flags needed before)

Edit casa_pipescript.py file (standard one/template:

https://science.nrao.edu/facilities/vla/data-processing/pipeline#the-casa-pipescript-py-file

And to run it:

- 1. > /path/to/casa/bin/casa --pipeline
- 2. On NRAO machines: > casa-pipe
- 3. Or within CASA: > execfile(`casa_pipescript.py')



Final notes

Considerations before running pipeline

- → Disk space needed 3-4x raw data size, even more if imaging
- → Compute time: 30min to ... a few days (weeks)
- → If the default pipeline is appropriate for your science (e.g. scpectral lines, rerunning pipeline, etc)

VLA CASA Calibration Pipeline information at: http://go.nrao.edu/vla-pipe

→ CASA Integrated Pipeline & Scripted Pipeline available

NRAO HelpDesk: https://help.nrao.edu/

Submit your ticket under the Pipeline Department.



www.nrao.edu science.nrao.edu public.nrao.edu

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

