

The VLA Pipeline Claire Chandler







- 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:
 - 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

- 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.

Real-time pipeline:

- 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 X-band and higher for priority A+B projects; 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 5.4.1, similar to current real-time pipeline
 - See the VLA CASA pipeline guide at http://go.nrao.edu/vla-casa-tut

- Scripted pipelines for CASA versions through 5.3.0 also available
 - Provides more flexibility in how to use the pipeline, options suitable for spectral line datasets, mixed correlator set-ups, multi-band observations, etc.
 - Working to incorporate these into the CASA integrated pipeline

Will the Pipeline work for you?

- The pipeline successfully completes on ~92% 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
 - Pipeline includes Hanning smoothing, RFI flagging, and weight calculations that may not be appropriate for (some) spectral line projects.
 - No polarization calibration (yet) but can use pipeline output as a starting point.*
 - Will probably work for data taken since May 2012, may work for earlier EVLA data, likely that extra flagging and editing may be needed in these cases

*CASA 5.4.1 and later, requires that you have used the correct polarization intents when setting up the observations



Pipeline Requirements

- "Standard" Stokes I science SB means:
 - 128 MHz spws (64 MHz for L-band; default setup), but may work on other set-ups as well
 - Can work for narrower BWs, depends on the strength of the calibrators
 - Heuristics currently make some assumptions about the strength of the calibrators, in particular, the delay calibrator
 - currently requires the SNR=3 limit on initial gain calibration per integration
 - Contains correctly labeled and complete scan intents
 - And also that the observations have been set up correctly!



Pipeline Requirements

- Correct observation set-up
 - Independent of whether you want to run the pipeline!
 - Remember: simple observing set-ups are always easier to calibrate
 - Do not skimp on calibration to spend more time on your target –
 you may end up not being able to calibrate the target data at all
 - Spending 3 minutes pointing could buy you more sensitivity than doubling the time on your target.



Pipeline Requirements

Scan intents

- The pipeline relies entirely on correct scan intents to be defined in each SB
- In order for the pipeline to run successfully on an SB it must contain, at minimum, scans with the following intents:
 - A flux density scale calibrator scan that observes one of the primary calibrators (3C48, 3C138, 3C147, or 3C286)* – this will also be used as the delay and bandpass calibrator if no bandpass or delay calibrator is defined
 - Complex gain calibrator scans

^{*}Some of these calibrators are variable (especially 3C48, 3C138), use 3C286 if possible

(Real-Time) Heuristics (I)

- Assuming requirements are met, the pipeline:
 - Loads the data (SDM-BDF → MS)
 - Hanning smooths*
 - 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.

(Real-Time) Heuristics (II)

- 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.
- 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*
- *May want to modify inputs and/or omit entirely for spectral line.



Flow chart (CASA pipeline)

load main ms, extract
info about contents
 (hifv_importdata)

Hanning smooth the data (hifv_hanning)

apply online flags, deterministic flags (hifv_flagdata) set flux cal models (hifv_vlasetjy)

derive prior cals (antenna posn corrections, opacity, rq gains, etc.)
(hifv_priorcals)

test delay & BP cals (normalized due to prior gaincal)
(hifv_testBPdcals)

Flow chart

flag bad deformatters/spws (hifv_flagbaddef)

run rflag on calibrated delay & BP cals (hifv_checkflag)

"semiFinal" delay & BP cals (not normalized, sp. index of BPcal not yet determined) (hifv_semiFinalBPdcals, hifv_checkflag, hifv_semiFinalBPdcals)

split out calibrators into calibrators.ms, do a test gaincal to determine short and long solints (hifv_solint)

make amp gain table for flux density bootstrapping, and do flux density bootstrapping; includes deriving sp. index of calibrators and re-inserting into main ms (hify fluxboot)

make final cal tables (delay, BP, amp, phase) – includes redoing split, fluxboot, etc. (hifv_finalcals)

Flow chart

apply final calibrations to main ms (hifv_applycals)

run rflag on all fields, including target (hifv_targetflag)

run statwt to set weights using rms per spw (hifv statwt)

make some diagnostic plots (hifv_plotsummary)

prepare list of images to make (hif_makeimlist)

make images of calibrators per spw (hif_makeimages)

science quality imaging (interactive)

- Pipeline weblog is created real-time (check while running)
- Diagnostic plots and tables for most stages.
- CASA log file by task or as a whole
- Calibrator images per spw.

Example used here, and in our Pipeline CASA Guide:

casa.nrao.edu/Data/EVLA/Pipeline/CASA5.4.1/html/

casa.nrao.edu/Data/EVLA/Pipeline/CASA5.4.1/html/





By

By Topic By Task

Project Code N/A

Observation Overview

Project	uid://evla/pdb/14411854
Principal Investigator	Prof. Dominik A. Riechers
Observation Start	2013-03-23 05:09:03 UTC
Observation End	2013-03-23 08:05:36 UTC

Pipeline Summary

Pipeline Version	42192 (Pipeline-CASA54-P2-B)
CASA Version	5.4.1-31 (environment)
Pipeline Start	2018-12-09 05:19:55 UTC
Execution Duration	1 day, 6:43:00

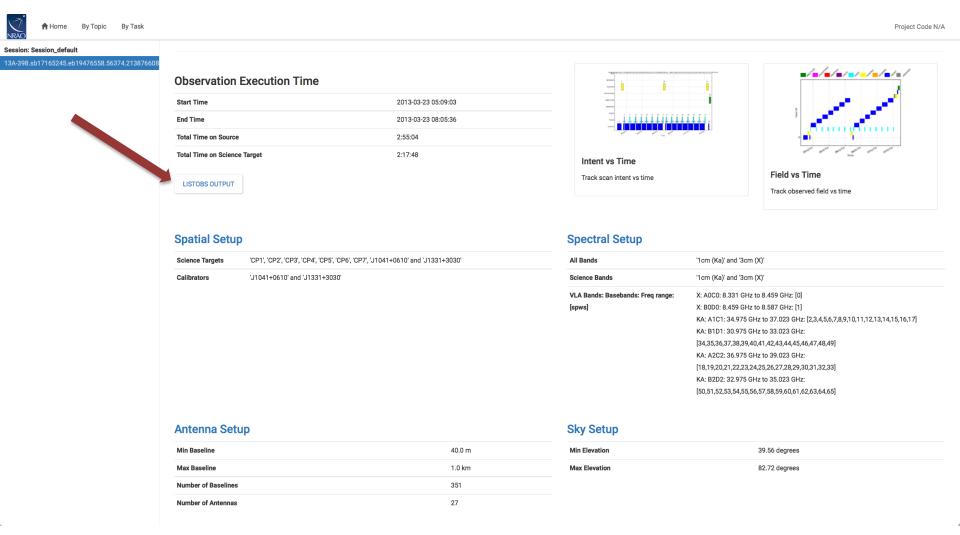
Observation Summary

				Time (UTC)			Baseline Length			
	Measurement Set	Receivers	Num Antennas	Start	End	On Source	Min	Max	RMS	Size
Scheduling Block ID: uid://evla/pdbsb/17165245										
Session: session_1										
	13A-398.sb17165245.eb19476558.56374.213876608796.ms	1cm (Ka) and 3cm (X)	27	2013-03-23 05:09:03	2013-03-23 08:05:36	2:17:48	40.0 m	1.0 km	441.9 m	146.0 GB











ZÍ

♠ Home

By Topic

By Task

Project Code N/A

BACK

Session: Session_default

13A-398.sb17165245.eb19476558.56374.2°

listobs.txt

```
MeasurementSet Name: /lustre/aoc/sciops/emomjian/pipe-demo/13A-398.sb17165245.eb19476558.56374.213876608796.ms
                                                MS Version 2
 Observer: Prof. Dominik A. Riechers
                Project: uid://evla/pdb/14411854
Observation: EVLA
Data records: 72213336
          Total elapsed time = 10767 seconds
 Observed from 23-Mar-2013/05:09:03.0 to 23-Mar-2013/08:08:30.0 (UTC)
 ObservationID = 0
           ArrayID = 0
     Timerange (UTC)
              Scan FldId FieldName
                               SpwIds Average Interval(s) ScanIntent
                            124254 [0,1] [1, 1] [OBSERVE_TARGET#UNSPECIFIED]
23-Mar-2013/05:09:03.0 - 05:12:00.0
                  0 J1041+0610
     05:12:03.0 - 05:16:57.0
                            2201472 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:17:02.0 - 05:17:05.0 3
                 1 J1041+0610
                             2106 [0,1] [1, 1] [UNSPECIFIED#UNSPECIFIED]
     05:17:06.0 - 05:20:30.0
                  1 J1041+0610
                            136188 [0,1] [1, 1] [CALIBRATE_POINTING#ON_SOURCE]
     05:20:33.0 - 05:21:57.0 5
                  0 J1041+0610
                            628992 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:22:03.0 - 05:23:00.0 6 2 J1041+0610
                            426816 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:23:03.0 - 05:29:27.0 7 3 CP1
                            2875392 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:29:30.0 - 05:32:24.0 8 3 CP1
                            1302912 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:32:30.0 - 05:33:54.0 9 2 J1041+0610
                            05:34:00.0 - 05:40:24.0 10 4 CP2
```

♠ Home

By Topic

By Task

Session: Session default

13A-398.sb17165245.eb19476558.56374.2°

listobs.txt

Project Code N/A

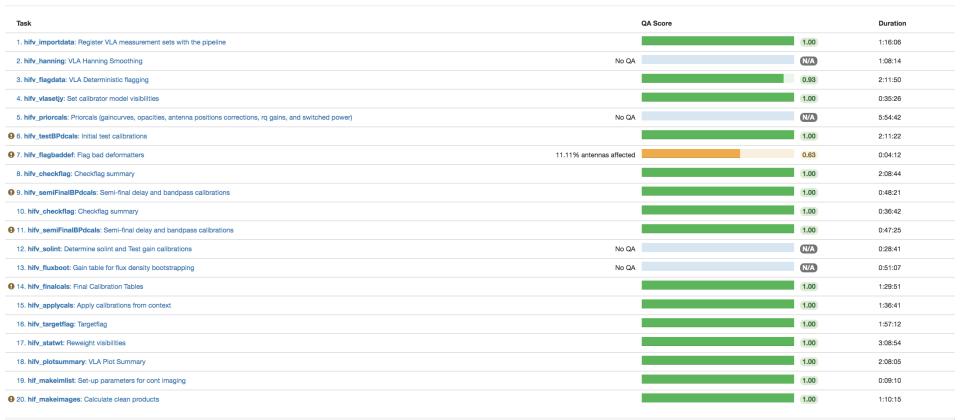
BACK

```
MeasurementSet Name: /lustre/aoc/sciops/emomjian/pipe-demo/13A-398.sb17165245.eb19476558.56374.213876608796.ms
                                                MS Version 2
 Observer: Prof. Dominik A. Riechers
                Project: uid://evla/pdb/14411854
Observation: EVLA
Data records: 72213336
          Total elapsed time = 10767 seconds
 Observed from 23-Mar-2013/05:09:03.0 to 23-Mar-2013/08:08:30.0 (UTC)
 ObservationID = 0
           ArrayID = 0
     Timerange (UTC)
              Scan FldId FieldName
                               SpwIds Average Interval(s) ScanIntent
                            124254 [0,1] [1, 1] [OBSERVE_TARGET#UNSPECIFIED]
23-Mar-2013/05:09:03.0 - 05:12:00.0
                  0 J1041+0610
     05:12:03.0 - 05:16:57.0
                            2201472 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:17:02.0 - 05:17:05.0 3
                 1 J1041+0610
                             2106 [0,1] [1, 1] [UNSPECIFIED#UNSPECIFIED]
     05:17:06.0 - 05:20:30.0
                  1 J1041+0610
                            136188 [0,1] [1, 1] [CALIBRATE_POINTING#ON_SOURCE]
     05:20:33.0 - 05:21:57.0 5
                  0 J1041+0610
                            628992 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:22:03.0 - 05:23:00.0 6 2 J1041+0610
                            426816 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:23:03.0 - 05:29:27.0 7 3 CP1
                            2875392 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:29:30.0 - 05:32:24.0 8 3 CP1
                            1302912 [2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,4
05:32:30.0 - 05:33:54.0 9 2 J1041+0610
                            05:34:00.0 - 05:40:24.0 10 4 CP2
```





Task Summaries



CASA logs and scripts

View, view in new tab or download casa-20181209-051926.log (35.2 MB)

 The following pipeline steps provide key checks for calibration quality:

hifv_flagdata
 deterministic flagged data fraction

hifv_testBPdcals hardware problems and other obs. issues

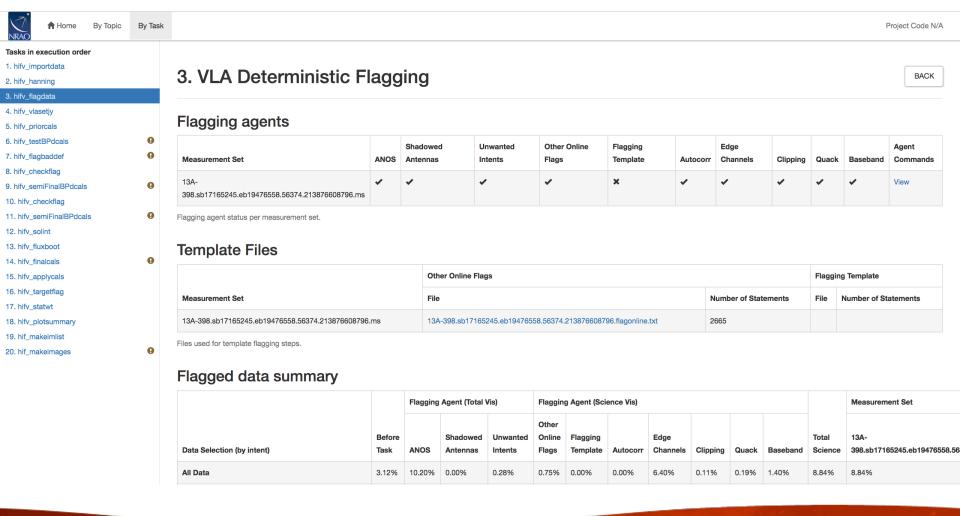
hifv_solint solution intervals for phase cals, input gain tables

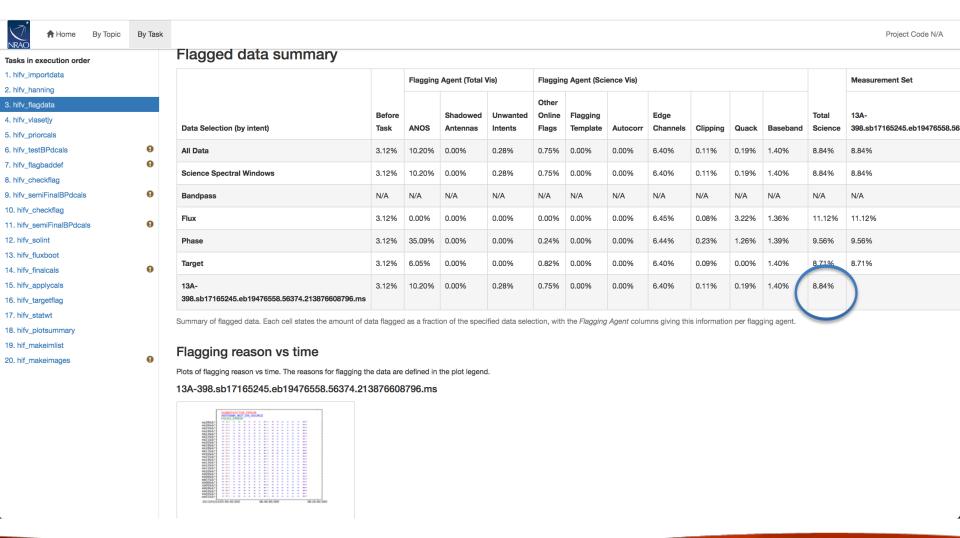
hifv_fluxboot fitted calibrator flux densities and spectral indices

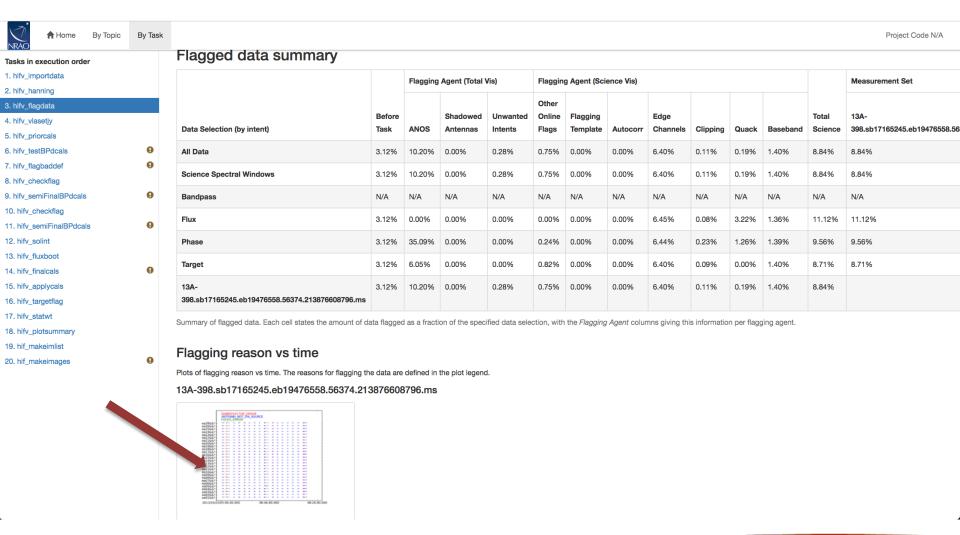
hifv_finalcals final calibration tables applied to the data

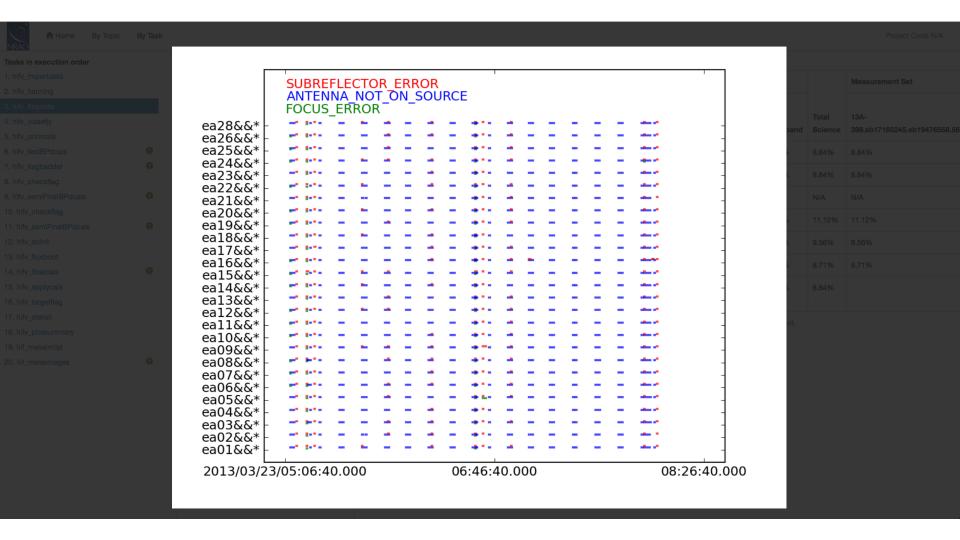
hifv_plotsummary useful diagnostic plots of calibrated data





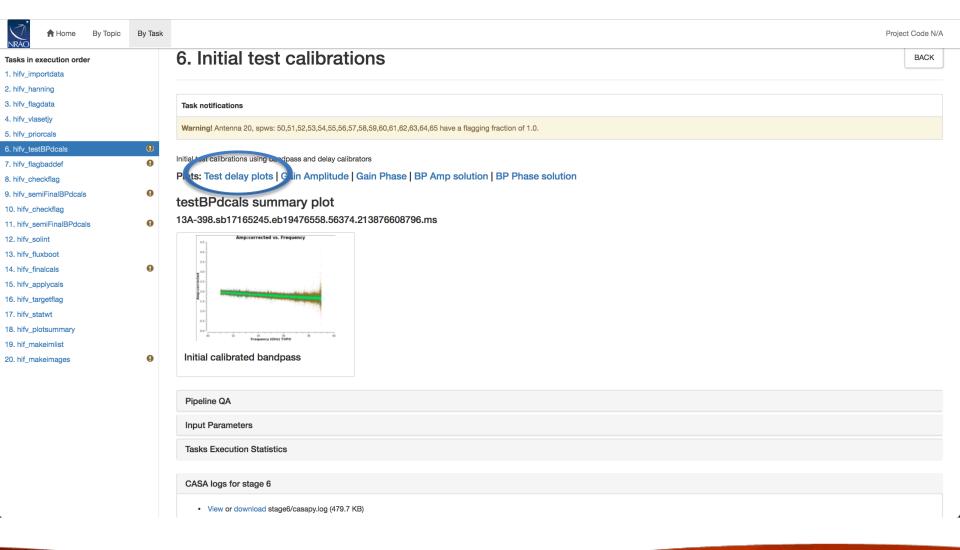




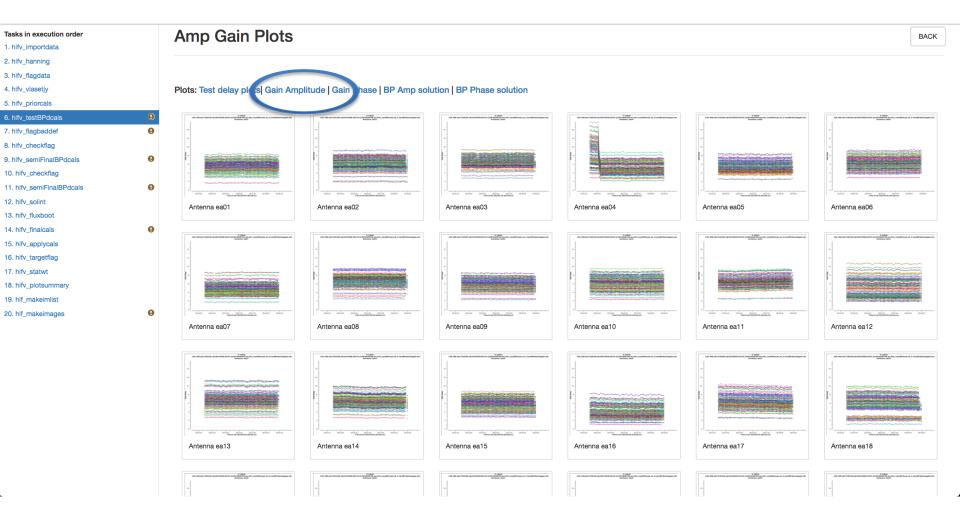








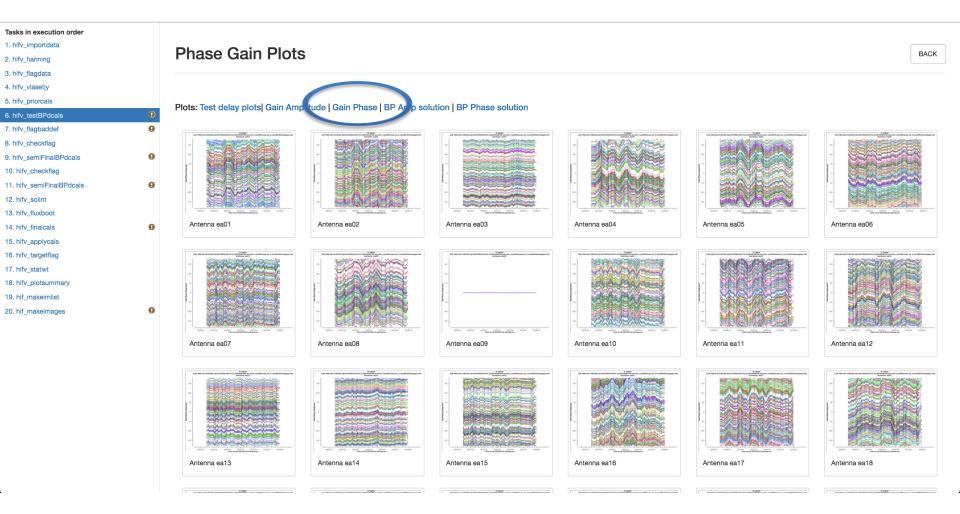






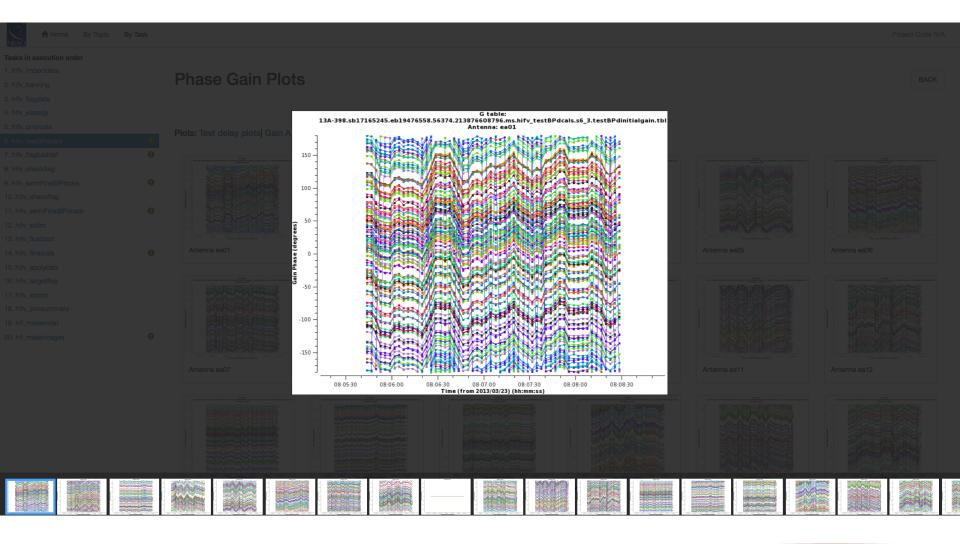






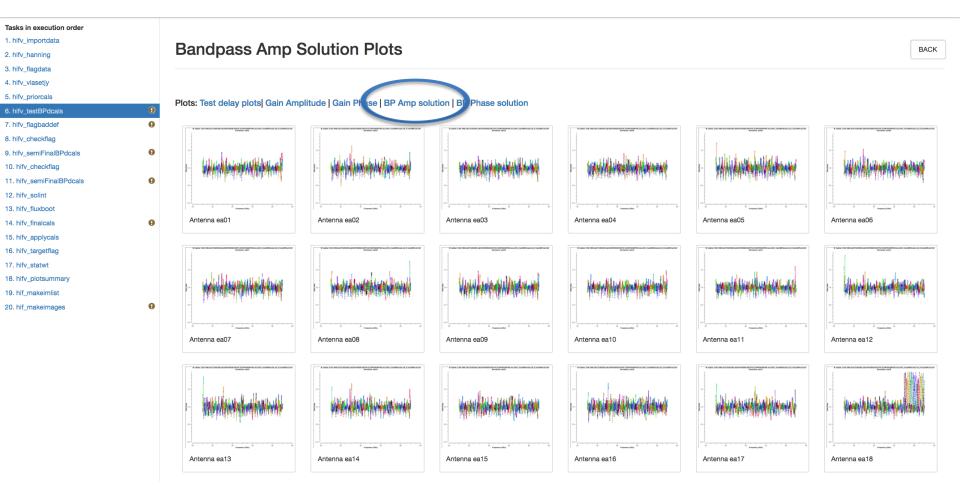








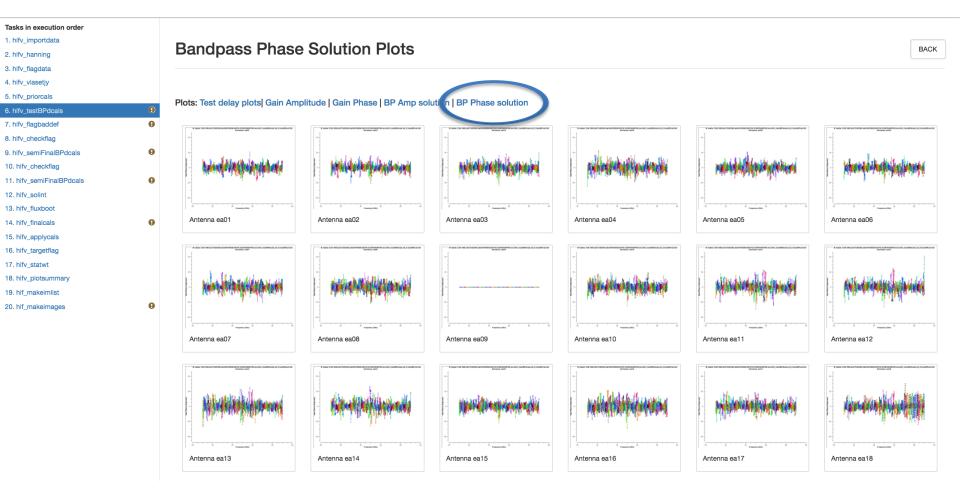








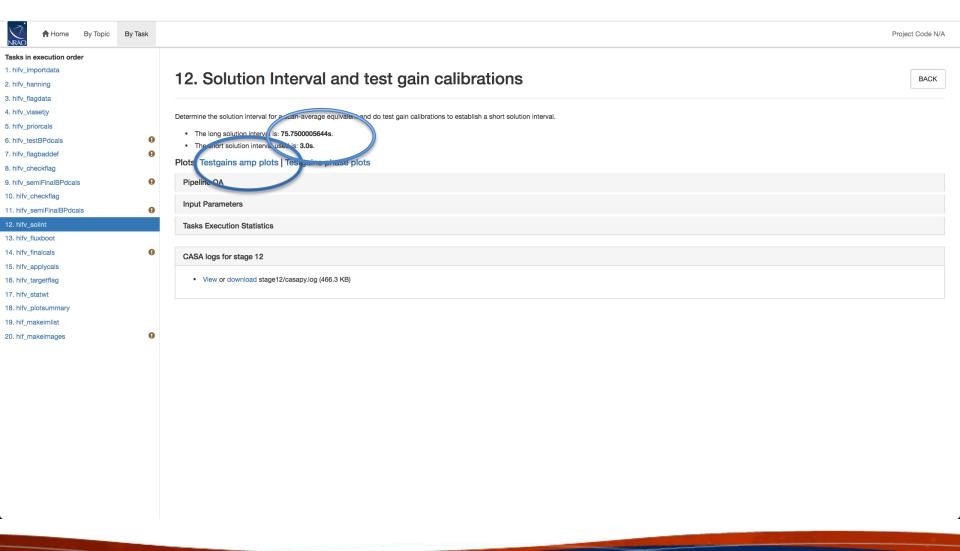






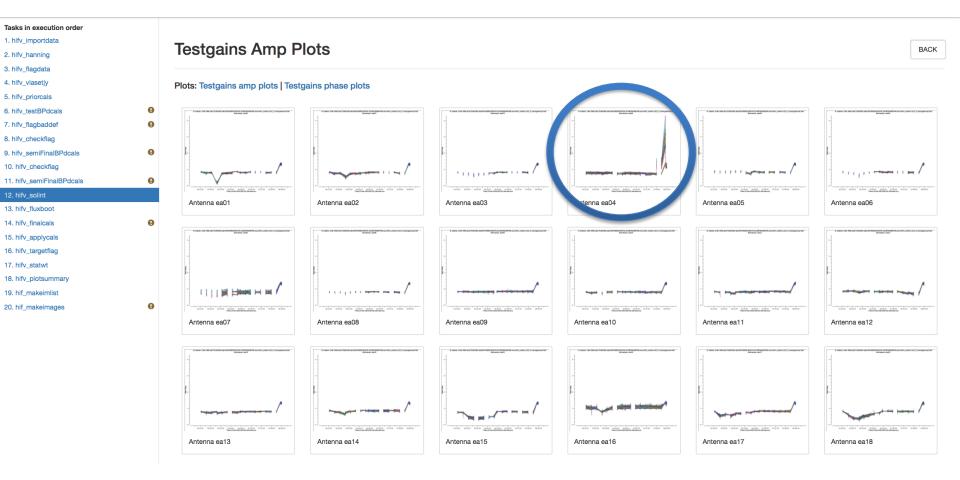


Gain Solution Intervals (hifv_solint)





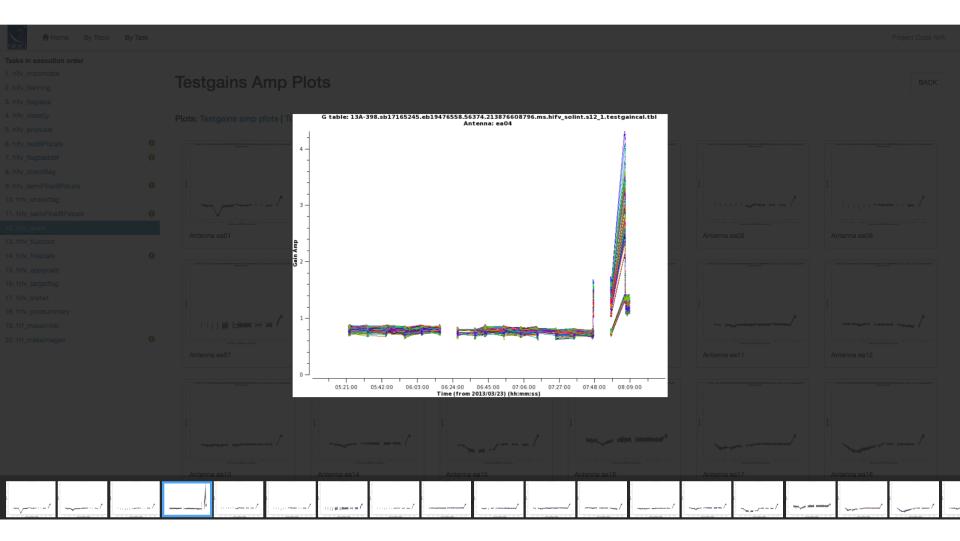
Gain Solution Intervals (hifv_solint)







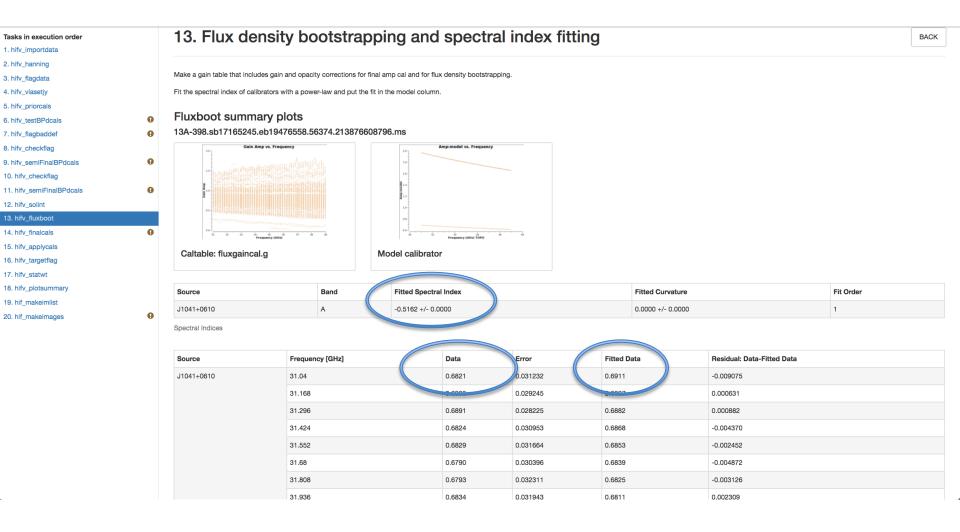
Gain Solution Intervals (hifv_solint)

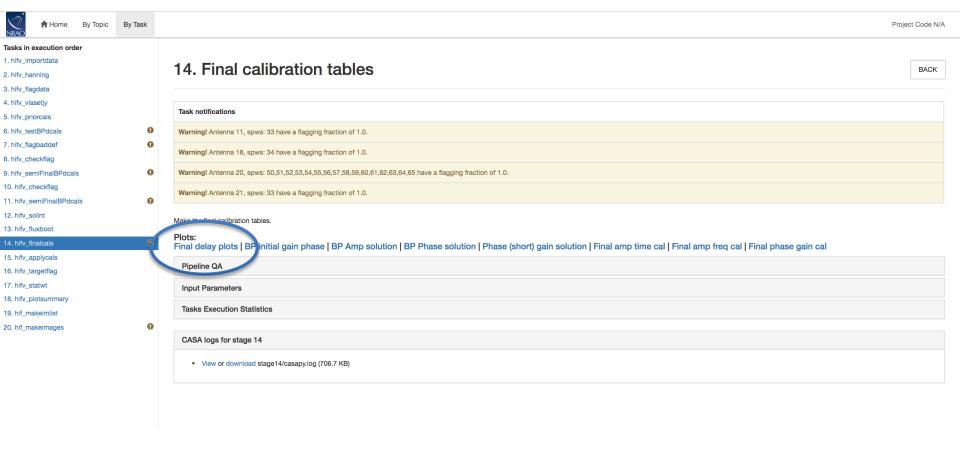




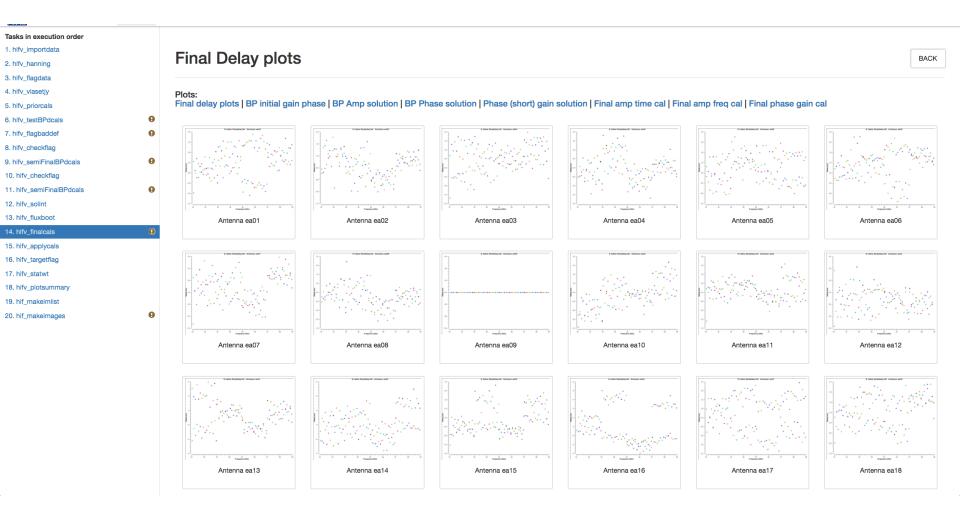


Flux Density Bootstrapping (hifv_fluxboot)

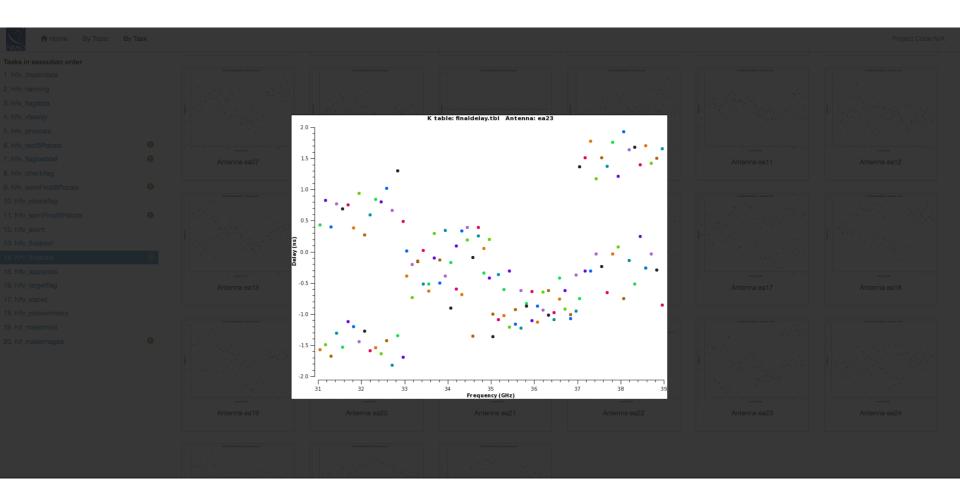




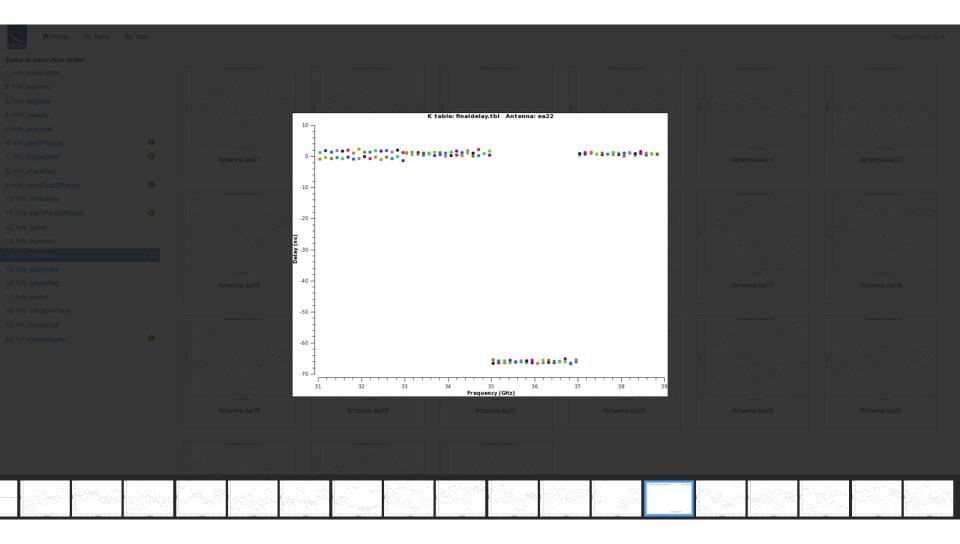




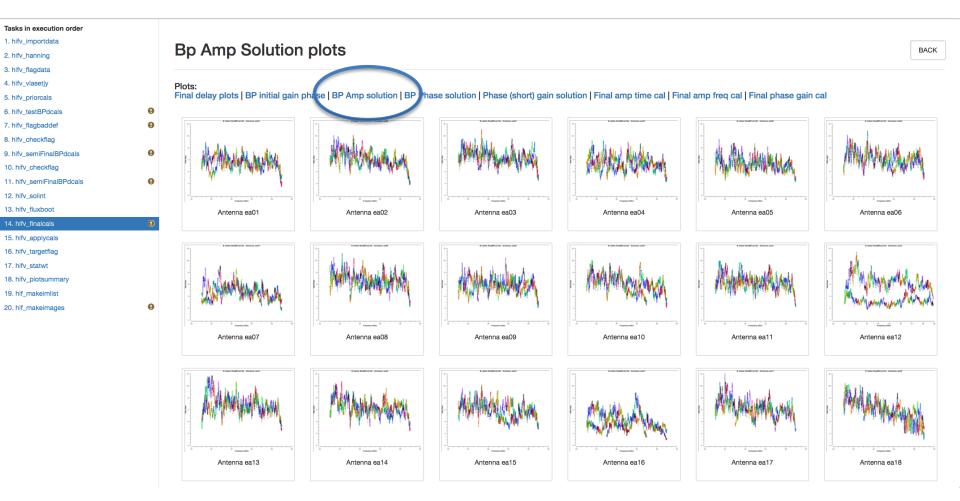






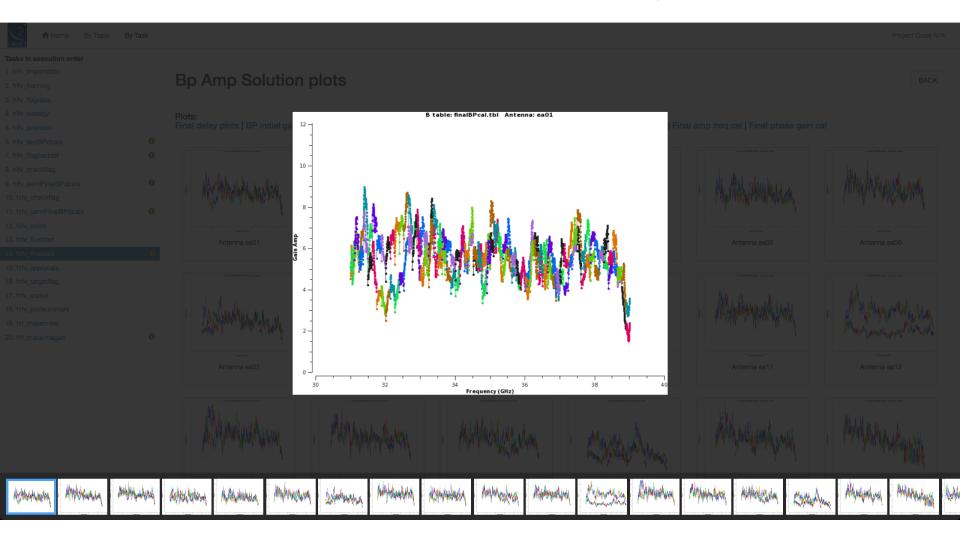






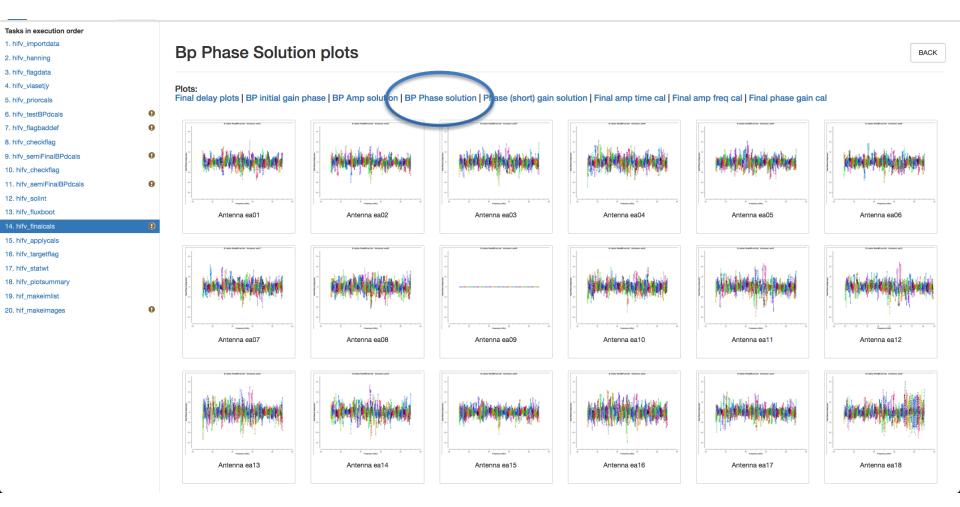


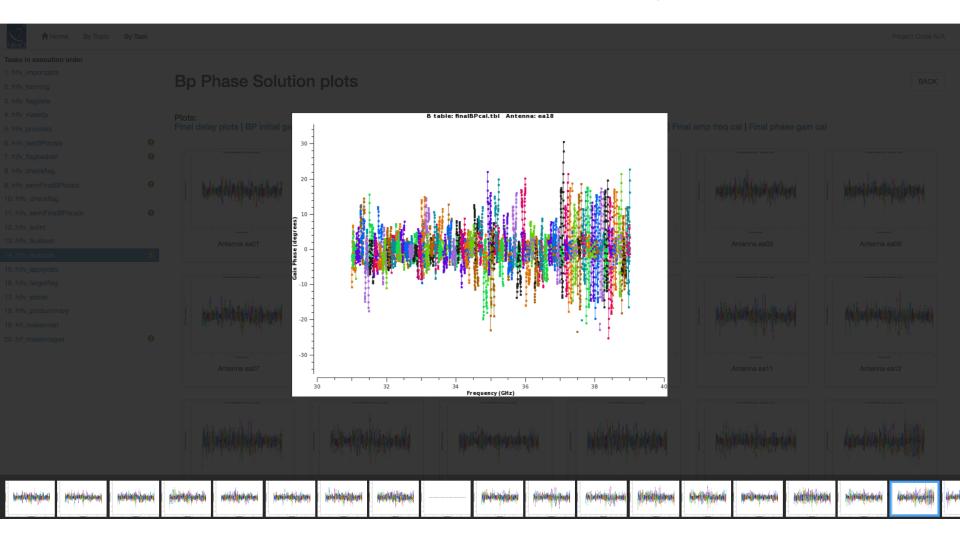






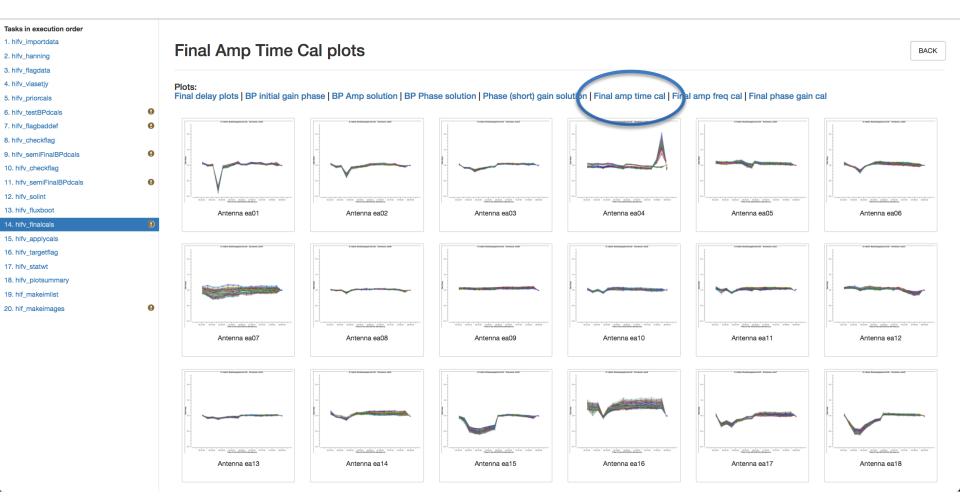




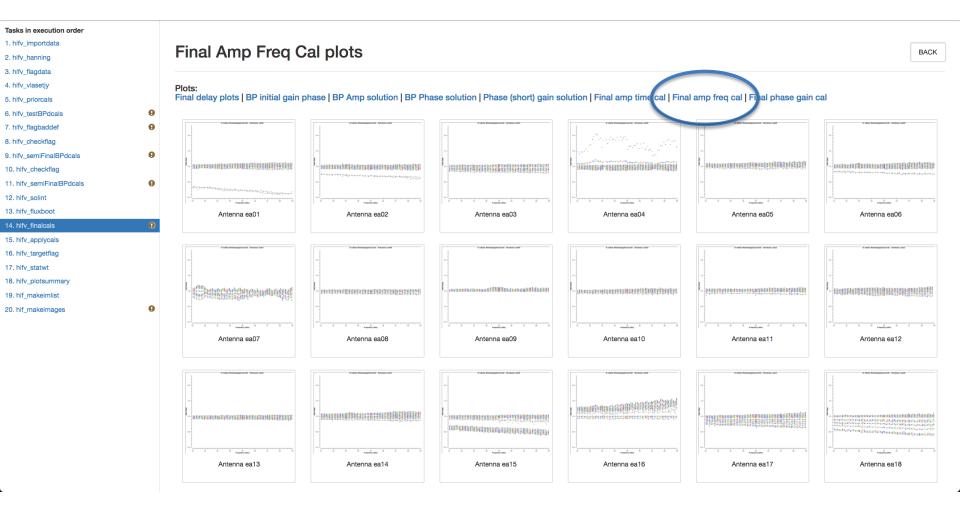




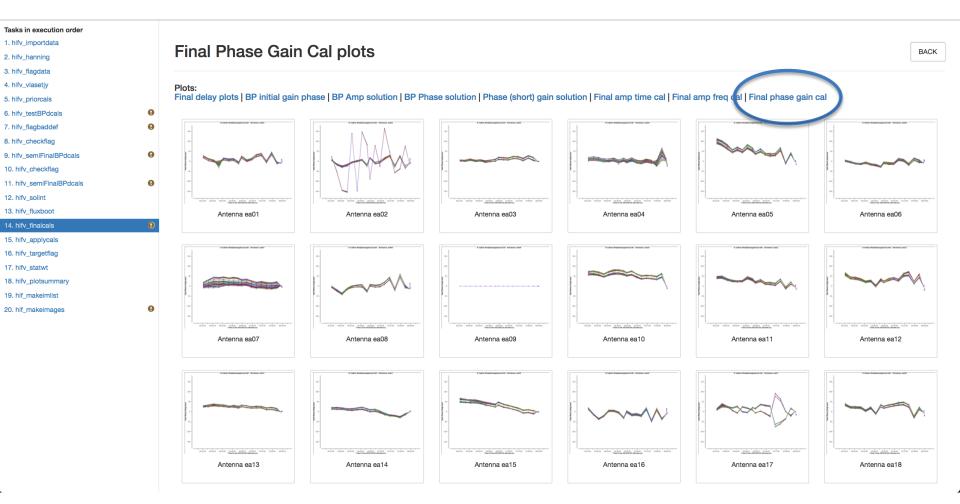
Final Cal Tables: amplitude and phase

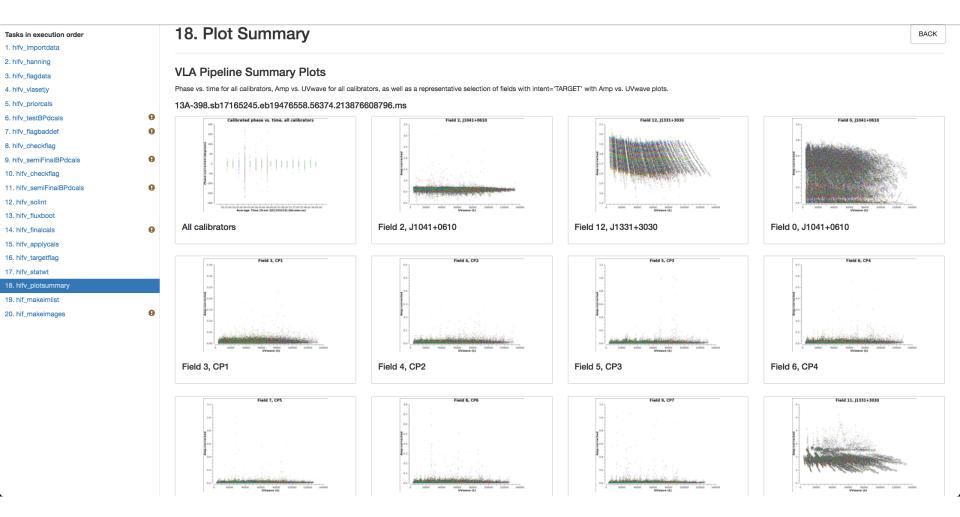


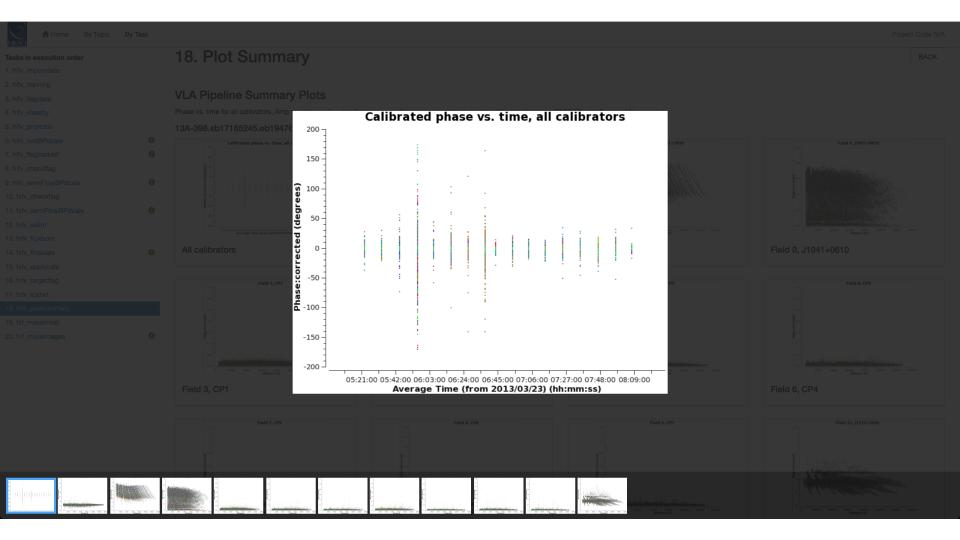
Final Cal Tables: amplitude and phase



Final Cal Tables: amplitude and phase

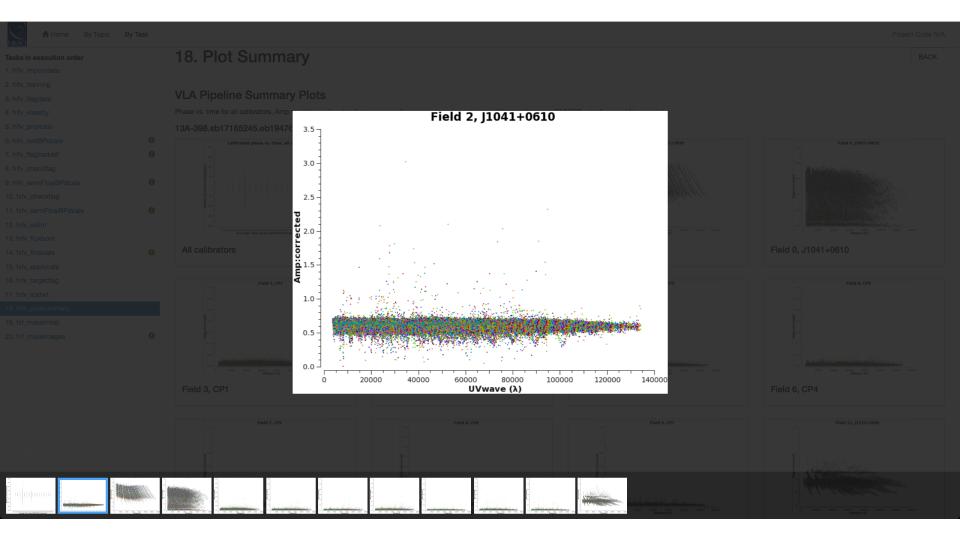






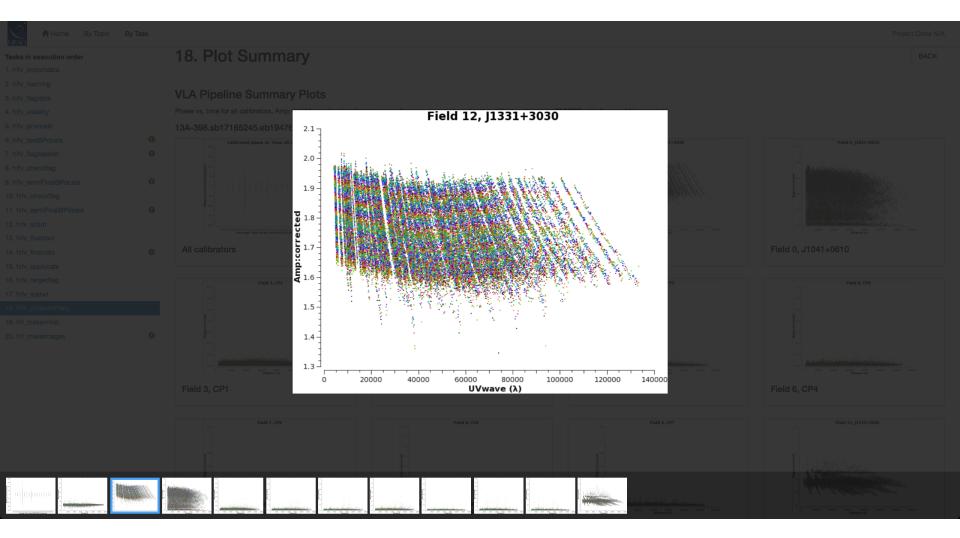






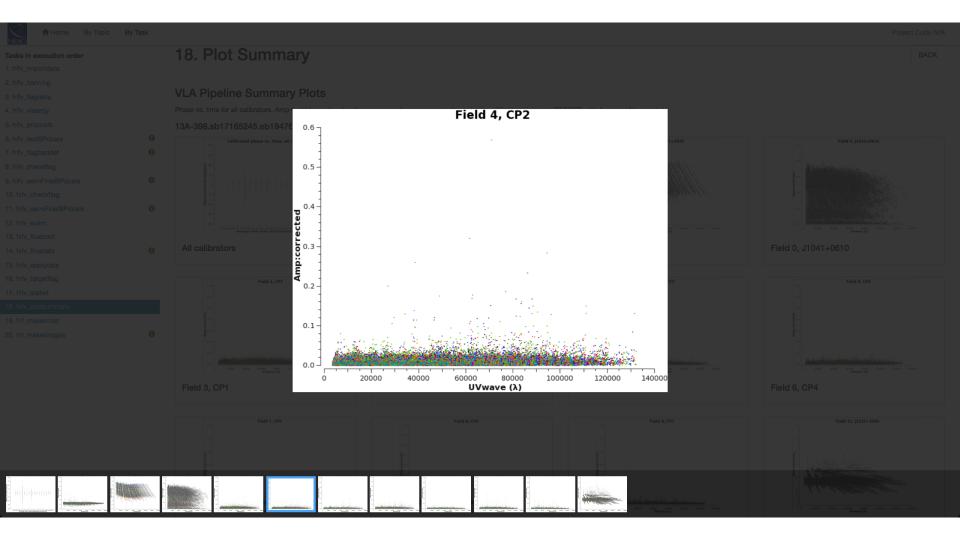








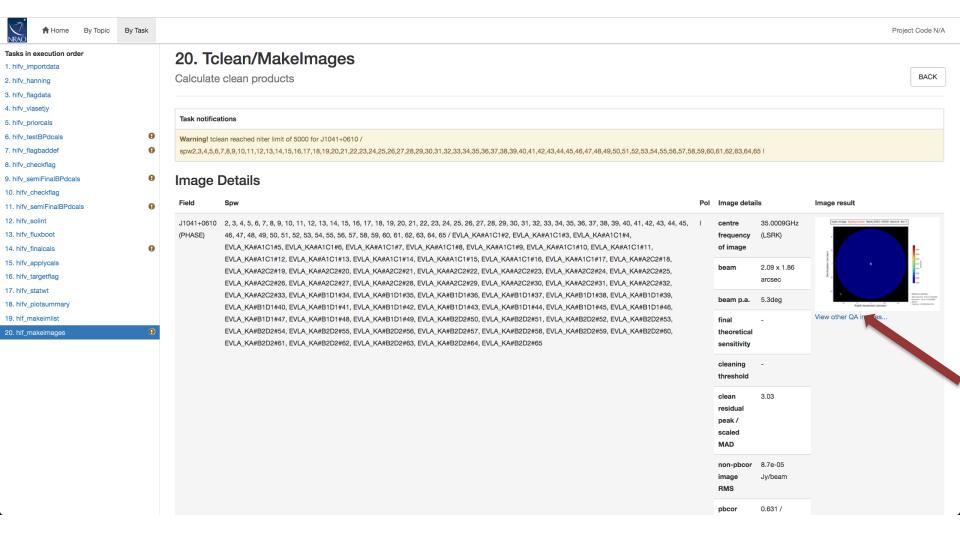




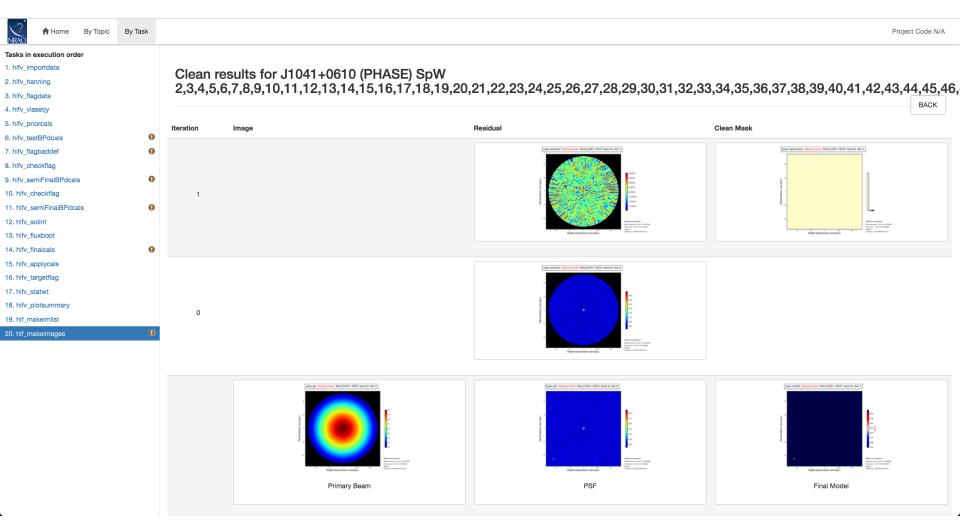




Calibrator Images (hif_makeimages)



Calibrator Images (hif_makeimages)



Pipeline Products and Outputs

- Flag versions and calibration tables (archived)
- Calibrated MS (available for 15 days, not archived)
- Logs, including weblog used by quality assurance (QA) staff and QA report.

Pipeline Products and Outputs

- The real-time pipeline produces a calibrated and flagged MS:
 - The products can be requested through the helpdesk (<u>help.nrao.edu</u>, VLA Pipeline Department):
 - For download over the internet, or for shipping on hard disk(s).
 - You may request a detailed QA2 report from the data analysts
 - If you are happy with the pipeline calibration, then:
 - Do further flagging if necessary
 - Split out your target and image
 - If you have the SDM or uncalibrated MS and the calibration and flag tables, instructions for applying flags and calibration tables may be found at http://go.nrao.edu/vla-pipe

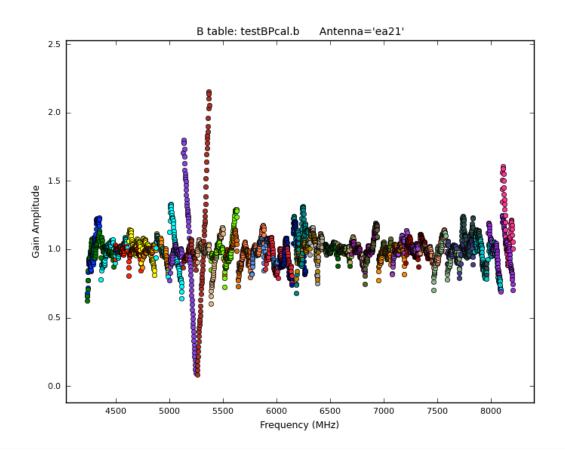
Pipeline Products and Outputs

- In some cases the pipeline and/or the MS may need to be modified
 - Download the SDM from the archive plus pipeline scripts
 - Follow the directions at http://go.nrao.edu/vla-pipe
- In some cases the pipeline heuristics may not be appropriate for your data (e.g., some L-band set-ups do not work well with the pipeline yet)
 - Reduce data by hand

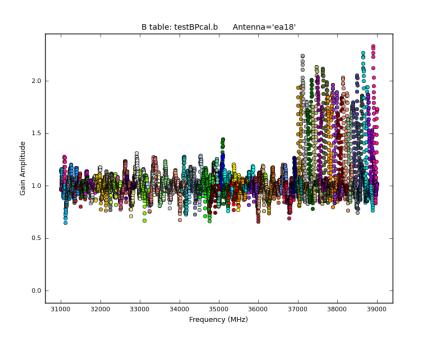


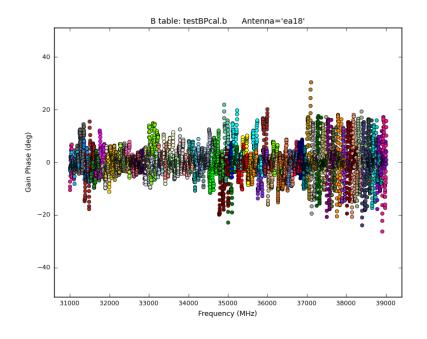
- In general the pipeline does very well, but there are possible failure modes:
 - No flux density or gain calibrator intents defined, or flux density calibrator not one for which we have models
 - work around in scripted pipeline
 - Wrong scan intents
 - modify Scan.xml in SDM; see https://science.nrao.edu/facilities/vla/dataprocessing/pipeline#section-28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

ea21 bandpass, bad data (DTS issue)

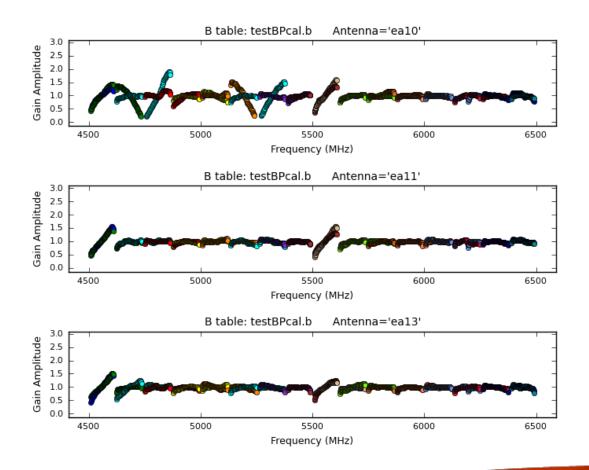


ea18 Amp and Phase affected (DTS issue for 37-39GHz)

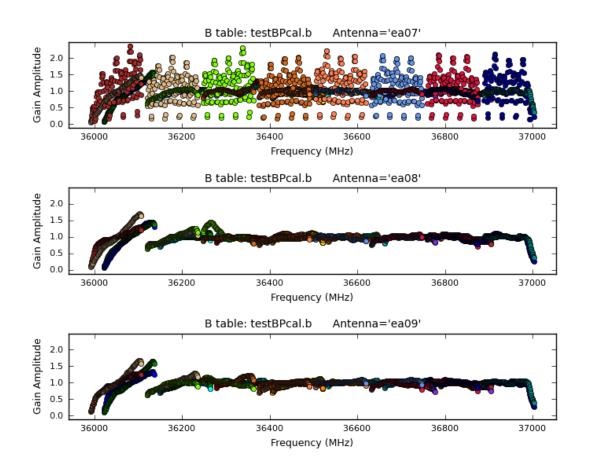




ea10 bandpass, bad data (DTS issue); ea11, ea12 OK



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



Spectral line data

- Several steps in the real-time pipeline may not be appropriate for spectral line data:
 - Hanning smoothing (increases effective channel width)
 - Flags 5% of *each* spw edge and the first and last 10 channels of each baseband
 - Last run of RFLAG on target (may eliminate your line as interference!)
 - Statwt calculates rms based on scatter of channels per spw, per visibility; may want to run manually with channel selection turned on to eliminate use of channels containing line emission in calculating the rms
 - ⇒ Specify a "cont.dat" file to avoid known lines for RFI flagging and statwt
- With the above modifications, the pipeline will work with spectral line data as long as the calibrators are strong enough

Mixed correlator set-ups

- With WIDAR capabilities it is common to observe both wide and narrow spws to obtain both continuum and spectral line data simultaneously or multiple receiver bands
 - A single heuristic (e.g., gain calibration solution interval) for entire dataset may not be appropriate

Solution:

- Run pipeline through application of deterministic flags, including Hanning smoothing if you are going to use it
- Split the MS by spw and/or scans
- Run pipeline on split MSs WITHOUT Hanning smoothing (you have already applied it, if you are going to use it)
- Warning: output flagging statistics may not be correct

Future Developments

- The NRAO Science Ready Data Products initiative is planning the following development over the next year:
 - Support for weak bandpass calibrators (implemented but not well tested)
 - Polarization calibration tested for VLA Sky Survey (S-band), need polarization calibrator models for other bands
 - More robust flux density bootstrapping that flags outliers
 - Improved RFI flagging and detection of system issues
 - Better support for multi-band observations
 - Use of switched power data for determining weights
 - Basic imaging pipeline



Questions?

VLA CASA Calibration Pipeline information at:

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

CASA Integrated Pipeline & Scripted Pipeline available

- Have Questions?
- Need Help?
- Report a bug?
- Use the NRAO HelpDesk: https://help.nrao.edu/
- Submit your ticket under the Pipeline Department.
- Please include specific details when submitting HelpDesk tickets.
 (Project code, SB number, CASA/PL versions, errors, etc.)



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.