EU ARC QA2 Training

D. Petry, ESO, 9-13 Feb 2015



Schedule

Monday (room Lupus (C3.01) 10:00-11:30 and 12:30-15:00):

- Introduction
- An example QA2 project
- Your first assignment

Tuesday: (room Lupus, 10:00-11:30)

- intermediate discussions on assignment (1)
- attempt to complete assignment (1) by end of day

Wednesday: (Pavo, 10:00-11:30)

- concluding discussion of assignment (1), introduction of assignment (2)
- The pipeline WebLog (Liz Humphreys) (11:00-11:30)
- work on assignment (2)

Thursday: (Lupus, 10:00-11:30)

- discussion on assignment (2), introduction of assignment (3)
- work on assignment (3)

Friday: (Lupus, 10:00-11:30)

- concluding discussion of assignment (3)
- packaging and AOB

Your offices and accounts:



C.3.03: Luke + Ilsang

B.3.48: Devaky + Nuria

C.2.23: Silvio + Ciro

We will form 2-person working groups

Group A: Paola - Matthieu

Group B: Luke - Ilsang

Group C: Silvio - Ciro

Group D: Gergely - Ruud

Group E: Devaky - Nuria

Cluster accounts: "almaXX", see the wiki page

Generally, log in via:

ssh -X almaXX@arcp1.hq.eso.org

qi (to get interactive shell on least loaded node)

see the cluster page and the cluster CASA session page





C.3.03: Luke + Ilsang

B.3.48: Devaky + Nuria

C.2.23: Silvio + Ciro

We will form 2-person working groups

Group A: Paola - Matthieu arcp4
Group B: Luke - Ilsang arcp5
Group C: Silvio - Ciro arcp6
Group D: Gergely - Ruud arcp7
Group E: Devaky - Nuria arcp8

Cluster accounts: "almaXX", see the wiki page

Generally, log in via the headnode arcp1:

ssh -X almaXX@arcp1.hq.eso.org

qi (to get interactive shell on least loaded node)

see the cluster page and the cluster CASA session page

This week we will use direct login to the "group node" $ssh - X \ almaXX@arcpN.hq.eso.org$, N = 4,5,6,7, or 8



Monday, room Lupus, 10:00 h – 11:30h and 12:30h-14:30h:

Outline

- The present project status w.r.t. QA2
- Overview of the QA2 procedure
- Brief mention of QA3
- Available sources of information for QA2 analysts
- The assignment process
- Organising your workspace
- The checklists
- A tour of the QA2 instructions wiki
- An example of a QA2 analysis
- Your first assignment

The present situation



Cycle 2 has gone well so far (numbers as of 31 January 2015)

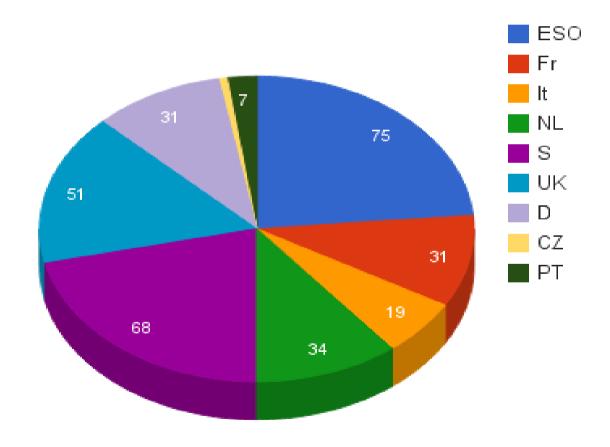
```
Total processed (including QA2_FAIL) since January 2013:
   238 (pass) + 80 (fail) = 318 SBs
   413 (pass) + 155 (fail) = 568 EBs
Total delivered: 238 SBs (413 EBs)
    Manual cal.: 196 SBs
  Pipeline cal.: 42 SBs
For the data observed between 2013/01/24 and 2014/08/27:
Overall Net analysis time median: 7 working days
Time from observation to delivery median 22 working days
(includes delays due to work at JAO)
Analysis effort 2013+2014 amounts to ca. 2000 working days
        (assuming 75% efficiency, this is ca. 7.5 FTE years)
Since Pipeline became available in October,
> 50% of the calibration is done by pipeline
```

The present situation



- many different analysts from all nodes, PT, and ESO participate

Cycle 1+2 EU ARC QA2 projects (as of 31/01/2015): processed SBs

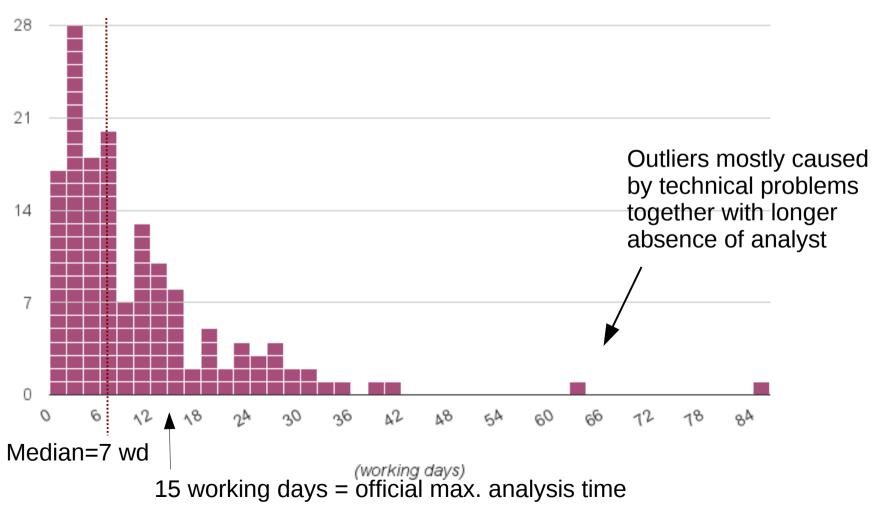


The present situation



- Net analysis time: range 1-84, median: 7 working days Record (held by M. Schmalzl): 1 day

Cycle 1+2 EU ARC QA2 projects (as of 1/09/2014): Net Analysis Time



The QA2 procedure in Cycle 2 (= the present status)



Leads: Balthasar Vila Vilaro + Eric Villard Analysts: ca. 8 JAO personel (Liza, Harold, Jose, ...)



The QA2 procedure in Cycle 2

What does the PI obtain from us?



```
-- 2013.1.00020.S/
   -- sq ouss id/
                                      Example: package for 1 EB
       -- group ouss id/
           -- member ouss id/
               -- README
               -- calibration/
                   -- uid A002 X86fcfa X14a6.calibration.tgz
                   -- uid A002 X86fcfa X14a6.ms.wvrgcal
                   -- uid A002 X86fcfa_X14a6.calibration.plots/
               -- log/
                   -- imaging.log
                   -- calibration.log
               -- product/
                   -- PKS1830 uid A002 X86fcfa X14a6.spw0 vel selfcal.flux.fits
                   -- PKS1830 uid A002 X86fcfa X14a6.spw0 vel selfcal.image.pbcor.fits
                   -- calibrated.cont.ms.image.continuum.source3.flux.fits
                   -- calibrated.cont.ms.image.continuum.source3.image.pbcor.fits
                   -- PKS1830 uid A002 X86fcfa X14a6.spw0 vel selfcal.mask/
                   -- calibrated.cont.ms.image.continuum.source3.mask/
               -- ga/
                   -- uid A002 X86fcfa X14a6 qa2 part1.png
                   -- uid A002 X86fcfa X14a6 qa2 part2.png
                   -- uid A002 X86fcfa X14a6 qa2 part3.png
                   -- uid A002 X86fcfa X14a6 textfile.txt
               -- script/
                   -- scriptForImaging.py
                   -- scriptForPI.py
                   -- uid A002 X86fcfa X14a6.ms.scriptForCalibration.py
```

The QA2 procedure in Cycle 2



Documented in:

- "How to Analyse ALMA Science Data" (Eric Villard)
- Our ARC Twiki Cycle2DataReduction
- The "QA2 Best Practices Document"

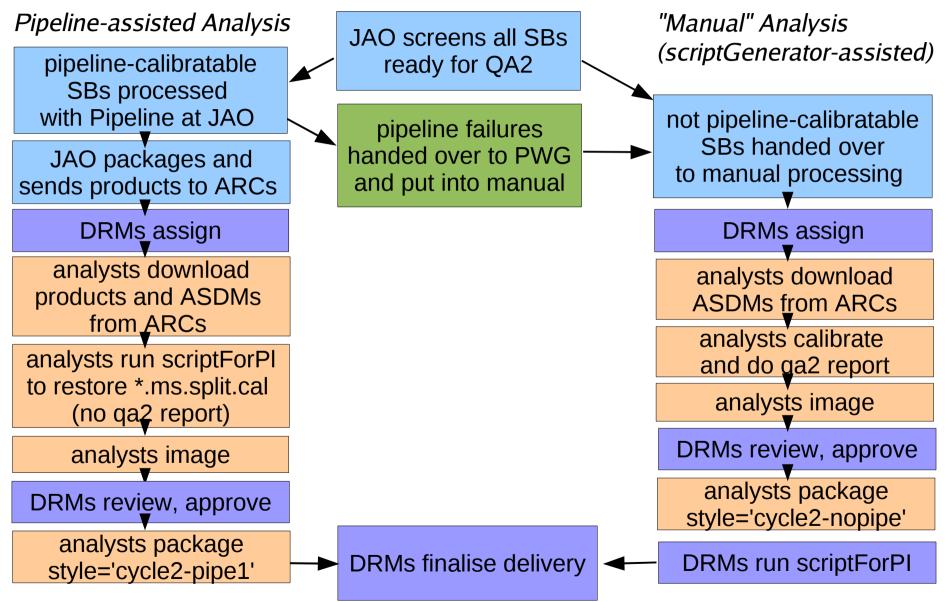
Basic steps:

- 1) When SB reaches sufficient number of executions, it is posted on an internal page and *DRMs distribute the work to analysts.*Bookkeeping is done via tickets SCOPS-390 (Cycle 1 carry-over) and SCOPS-1191 (Cycle 2).
- 2) Analyst at ARC performs QA2 as prescribed in above documents using specialised CASA scripts (the "script generator"). Duration: ca. 1 *day per ASDM*
- 3) When finished, analyst hands over products to DRM
- 4) When approved by DRM, data and products are delivered to PIs

ARC node analysts may carry out the QA2 work at their home institute. (Some SBs are also reduced at JAO. Here the leads play the role of the DRM.)

QA2 Workflow





The QA2 procedure in Cycle 2



Timeline for a QA2 procedure at the EU ARC:

Workflow A: Calibration and imaging at the ARC

- 1) DRMs finds SB ready for QA2
- 2) DRM finds volunteer analyst (ideally CS, not PI; otherwise back to JAO)
- 3) analyst starts QA2 data processing (data is in ESO archive copy)
- 4) analyst gives preliminary feedback about data quality after one week at the latest
- 5) analyst completes analysis including imaging and QA2 report after three weeks at the latest (15 working days)
- 6) analyst posts qa2 diagnostic results (ticket.zip) and checklists on JIRA reduction ticket for DRM to approve
- 7) upon approval: analyst packages and puts on server, DRM downloads after four weeks at the latest (20 working days)
- 8) DRM verifies and if necessary corrects packaging, tests scriptForPI
- 9) DRM hands finished tar ball to Archive Operations Group (AOG)
- 10) AOG uploads tarball to JAO and waits for notification that ingest is fine
- 11) AOG sends delivery notification to PI.

The QA2 procedure in Cycle 2



Timeline for a QA2 procedure at the EU ARC:

Workflow B: Calibration at JAO, imaging at the ARC

- 1) JAOs finds SB ready for QA2 and runs pipeline on it
- 2) JAO packages pipeline products and puts them on JAO server
- 3) DRM transfers package to ESO and finds volunteer analyst
- 4) analyst gets data and package from ESO; runs scriptForPI to get MS
- 5) analyst runs checkRMS and gives preliminary feedback after 3 working days at the latest
- 6) analyst completes imaging after two weeks at the latest (10 working days)
- 7) analyst posts qa2 diagnostic results (ticket.zip) and checklists on JIRA reduction ticket for DRM to approve
- 8) upon approval: analyst packages and puts on server, DRM downloads after three weeks at the latest (15 working days)
- 9) DRM verifies and if necessary corrects packaging, tests scriptForPI
- 10) DRM hands finished tar ball to Archive Operations Group (AOG)
- 11) AOG uploads tarball to JAO and waits for notification that ingest is fine
- 12) AOG sends delivery notification to PI.

QA2 at the EU ARC



Infrastructure at the EU ARC:

- The ARC cluster 8 nodes: 12 x Intel Xeon 2.4 GHz, 32 GB RAM, 120 TB Lustre filesystem (only 70% usable: 84 TB actual space)

Performance ≈ powerful personal workstation (supports up to 6 typical CASA instances at full speed)

Software:

- We are about to switch to CASA 4.3 but still using CASA 4.2.2
- AIV analysisUtils and related modules in the JAO CVS repository accessible via our mirror in the EU ARC svn repository http://svnhq16.hq.eso.org/p16/eso/arc/software/jao-mirror/AIV/science/

Concerning the role of script generator and pipeline



The analysisUtils script generator developed by Eric Villard et al. is essential to cross-check and verify the pipeline.

The analysisUtils pipeline will co-exist with the pipeline in the long run. Will serve as a *test-bed* for new analysis procedures and help with problematic/unusual data sets.

QA3 in Cycle 2



1.Initialization of Process

PI or CS on PI's behalf reports problem to the ARC via helpdesk

2.Initial Review

Helpdesk ticket assigned to DRM who finds a reviewer for it. JAO will maintain QA3 wiki to collect issues and solutions.

3.QA3 JIRA Ticket Creation

If the solution to the problem is not yet known, a SCOPS ticket is filed. All DRMs monitor the further progress of the ticket and inform CSs.

4.Investigation of the Issue

Remaining undelivered data of the project put on hold. JAO forms team to solve the issue.

5. Actions derived from Solution/or lack thereof

Possibilities: reprocess, reobserve + process,

Reobserve+process *other* projects; worst case: reobservation impossible

6.Extensions of the Proprietary Periods

If PI reports within 2 months: full extension; otherwise: only by QA3 time.

Summary: Available sources of information



... are listed at the top of the instructions wiki

http://www.eso.org/projects/alma/arc/tw/bin/viewauth/Offline/Cycle2DataReduction

under

"Relevant Documents"

- the "How-To"
- the "Best Practices"

For PIs there is now also the document "ALMA QA2 Data Products" (available on the ALMA Science Portal), Cycle 2 version still in the works

and

"Reporting and News etc."

- the JIRA tickets to track procedures
- the JIRA tickets to track project preparation
- the JIRA tickets to track project data reduction

The assignment process



- QA2 analyst is trained
- Becomes member of euarc-qa2@eso.org list
- DRM posts Doodle poll of monthly availablity
- Analyst enters availability in monthly **Doodle poll** (should only enter if available at least for 5 consecutive days 50% full time)
- When SBs need analysis, DRM sends email to available people from adequate node (if possible where contact scientist is based)
- Analyst confirms availability
- DRM sends assignment email

The assignment email



```
Hi, thanks for agreeing to perform QA2 on project 2012.1.00994.S SB(s) SPT_0125-50 http://jira.alma.cl/browse/SCOPS-574 project preparation ticket http://jira.alma.cl/browse/SCOPS-739 data reduction ticket The EB UIDs are uid___A002_X78774a_X5e5
```

The data can be downloaded at http://almascience.eso.org/arcdistribution/qa2-for-nodes/92e05474ec6959ecdab324510fde34d

The md5 checksum of each tar file can be found in the file md5sums.txt . In case of doubt verify by running "md5sum" on your downloaded copy.

Please read the instructions at http://www.eso.org/projects/alma/arc/tw/bin/view/Offline/Cycle1DataReduction

and don't hesitate to ask in case of any problems (ideally by replying to this email in order to stay in the same thread).

Please pay special attention to the new EB and SB checklists (one EB checklist for each EB, one SB checklist for each SB) which should at the end of your work be attached separately to the data reduction ticket. See e.g. proj 882, http://jira.alma.cl/browse/SCOPS-655.

Organising your workspace



... is described in the instructions wiki

http://www.eso.org/projects/alma/arc/tw/bin/viewauth/Offline/Cycle2DataReduction

under

"Setting up the directory structure"

Basically:

projectX SBY/

science/ -- your snapshot of the analysis utils

analysis/ -- the later packaging path

calibrated/ -- where the *.ms.split.cal MS from each EB should be moved

and where flux calibration and imaging take place

X123/ --

X456/ --

×789/ -- where the calibration for each EB takes place

At ESO, use the *startqa2.sh* script.

The EB Checklist (only used in manual calibration)



For each EB, you should track your analysis in the EB checklist

All checklist forms can be found in http://svnhq16.hq.eso.org/p16/eso/arc/software/jao-mirror/AIV/science/qa2/checklists

For Cycle 2 use: Cycle2_QA2_Checklist.txt

Rename to Checklist-EB-UID.txt (unique name)

To be attached to the Data Reduction JIRA Ticket.

http://jira.alma.cl/secure/attachment/180928/Checklist-uid___A002_X86fcfa_X14a6.txt

An EB Checklist example (header)

```
EUROPEAN ARC
ALMA Regional Centre
```

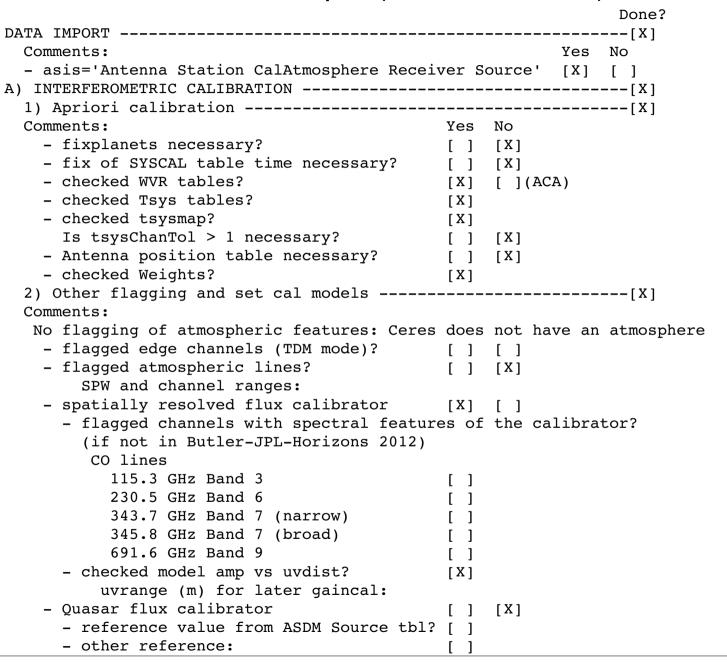
****** ALMA Cycle 2 OA2 Calibration Checklist ********** Checklist Version: \$Id: Cycle2 QA2 Checklist.txt,v 1.1 2014/07/11 12:08:18 dpetry Exp \$ CASA version: 4.2.1 analysisUtils version: \$Id: analysisUtils.py,v 1.1872 2014/08/16 20:30:48 thunter Exp \$ Reduction Start Date : 2014/08/18 Reduction Completion Date: 2014/08/20 Analyst Name(s) : D. Petry Project code : 2013.1.00020.S Contact Scientist: Ivan Marti-Vidal PΙ : S. Muller : PKS1830-211 b 07 TE SB Name : uid A002 X86fcfa X14a6 ASDM UID Observation Dates and Total Integration Time (from listobs): Data records: 865920 Total integration time = 2952.1 seconds Observed from 18-Jul-2014/03:26:09.1 to 18-Jul-2014/04:15:21.2 (UTC) : 0.5 mm PWV Number of good antennas: 29 (flagged: DA54 bad Tsys, DV07 bad WVR (may be rescued), DV14 bandpass) Band(s) : 7 TDM/FDM/both : FDM bandwidth switching (different spectral setup for phasecal and target): yes/no? no Fields by intent (paste from reduc script): AMPLI, ATMOSPHERE, BANDPASS, FLUX, FOCUS, PHASE, POINTING, TARGET # CALIBRATE AMPLI: J1924-292 # CALIBRATE ATMOSPHERE: J1924-2914, J1924-292, PKS1830-211 # CALIBRATE BANDPASS: J1924-2914 # CALIBRATE FLUX: J1924-292 # CALIBRATE PHASE: J1832-2039 # CALIBRATE POINTING: J1924-2914 # OBSERVE TARGET: PKS1830-211

An EB Checklist example (header)

```
EUROPEAN ARC
ALMA Regional Centre
```

```
****** ALMA Cycle 1 OA2 Calibration Checklist *********
Checklist Version: $Id: Cycle1 QA2 Checklist.txt,v 1.8 2014/01/13 14:30:33 dpetry Exp $
CASA version: 4.2.1
analysisUtils version: $Id: analysisUtils.py,v 1.1523 2014/04/12 21:38:53 thunter Exp $
Reduction Start Date
                        : 13 Apr 2014
Reduction Completion Date: 16 Apr 2014
Analyst Name(s) : M Schmalzl
Project code
                : 2012.1.00350.S
Contact Scientist: K Wang
                : C Caceres
PI
                : Continuum + CO line - Chamaleonis
SB Name
                : uid A002 X762210 Xf0
ASDM UID
Observation Dates and Total Integration Time (from listobs):
Total integration time = 2162.3 seconds
               18-Mar-2014/09:13:01.6 to 18-Mar-2014/09:49:03.9 (UTC)
Observed from
PWV
                      : 2.361 mm
Number of good antennas: 24
Band(s)
                      : 6
TDM/FDM/both
                      : FDM
bandwidth switching (different spectral setup for target and phasecal): yes/no? no
Fields by intent (paste from reduc script):
AMPLI, ATMOSPHERE, BANDPASS, FLUX, FOCUS, PHASE, POINTING, TARGET
# CALIBRATE AMPLI: Ceres
# CALIBRATE ATMOSPHERE: Ceres, J0635-7516, J1107-4449
# CALIBRATE BANDPASS: J1107-4449
# CALIBRATE FLUX: Ceres
# CALIBRATE PHASE: J0635-7516
# CALIBRATE POINTING: J0635-7516, J1107-4449
# OBSERVE TARGET: CHX 22,V star EH Cha,V star EI Cha,V star HI Cha
```

An EB Checklist example (sections A 1, 2)





An EB Checklist example (sections A 3)



Done?

3) bandpass -----[X] Comments:

The automatically chosen antenna DV11 showed some unwanted effects (low gain in some scans, strong spike in the bandpass). I changed it to DA47.

- Is REFANT chosen by script gen. good? [] [X]
 Name of REFANT: DA47
- checked .ap_pre_bandpass table? [X]
- mixed: used solint = 'inf,XMhz'? []
- TDM and mixed: checked .bandpass table? []
- FDM: checked .bandpass_smooth20ch table? [X]

An EB Checklist example (sections A 4)



Done?

4) amp cal of all calibrators (gaincal) -----[X]

The derived flux densities for J1107-4449 and J0635-7516 are higher than what is returned by aU.getALMAFluxForMS. At 221.418 GHz, I get the following values: aU: 0.85 Jy, fluxscale: 0.98 Jy J1107 aU: 0.57 Jy, fluxscale: 0.63 Jy J0635 The fluxscale values are 1.15 and 1.09 times higher. The time difference between the observations and the values from getALMAFluxForMS are 3 days. - For resolved flux cal: limited uvrange? [] [X] - For resolved flux cal: paste .fluxscale file here Flux density for J1107-4449 in SpW=0 is: 1.0093 + /- 0.00885402 (SNR = 113.993, N = 20) Flux density for J1107-4449 in SpW=1 is: 1.00123 +/- 0.00922962 (SNR = 108.48, N = 20) Flux density for J1107-4449 in SpW=2 is: 0.958836 + /- 0.0097336 (SNR = 98.5079, N = 20) Flux density for J1107-4449 in SpW=3 is: 0.948837 +/- 0.0104096 (SNR = 91.1501, N = 20) Flux density for J0635-7516 in SpW=0 is: 0.653922 + /- 0.00668452 (SNR = 97.8265, N = 18) Flux density for J0635-7516 in SpW=1 is: 0.648357 + -0.00640269 (SNR = 101.263, N = 18) Flux density for J0635-7516 in SpW=2 is: 0.604962 + - 0.00730067 (SNR = 82.8638, N = 18) Flux density for J0635-7516 in SpW=3 is: 0.599123 +/- 0.00746842 (SNR = 80.2209, N = 18) Fitted spectrum for J1107-4449 with fitorder=1: Flux density = 0.97925 + - 0.000751038(freg=221.418 GHz) spidx=-0.824988 +/- 0.0238147 Fitted spectrum for J0635-7516 with fitorder=1: Flux density = 0.626134 +/- 0.000527653(freq=221.418 GHz) spidx=-1.22514 +/- 0.0262835 - checked .phase int table? [X] - checked .ampli inf table? [X] - checked .phase inf table? [X] - checked uv data on calibrators? [X]

An EB Checklist example (sections A 5, 6, 7, B)



Not anonymous FTP! For user and password, see QA2 Instruction wiki section "In the case your data was of good quality: upload flux information!"

The SB Checklist (always used)



For each SB, you should track your analysis steps *after* basic calibration in the SB checklist

All checklist forms can be found in http://svnhq16.hq.eso.org/p16/eso/arc/software/jao-mirror/AIV/science/qa2/checklists

For Cycle 2 use: Cycle2_QA2_SB_Checklist.txt

Rename to Cycle2_QA2_SB_Checklist_SB-name.txt

To be attached to the Data Reduction JIRA ticket.

Examples abound on JIRA.

An example is at

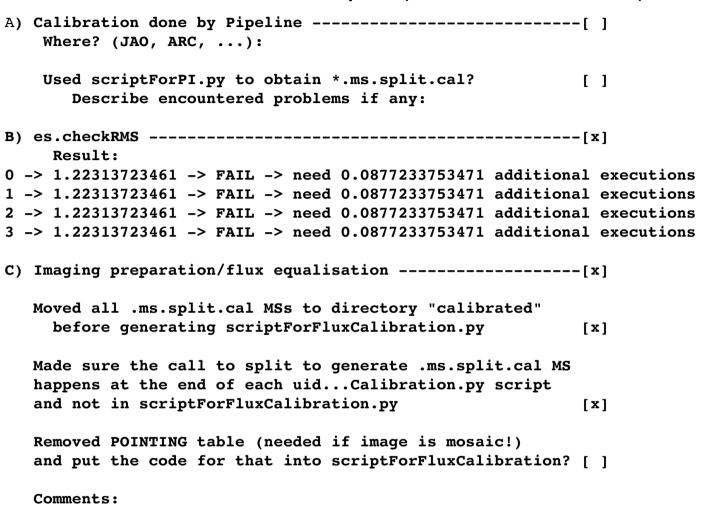
http://jira.alma.cl/secure/attachment/206252/Cycle2-SB-checklist-Orion_In_a_03_TE.txt

An SB Checklist example (header)



****** ALMA Cycle 2 QA2 SB Analysis Checklist *************** Checklist Version: \$Id: Cycle2 QA2 SB Checklist.txt,v 1.2 2014/08/13 08:15:13 dpetry Exp \$ CASA version: 4.2.1 analysisUtils version: \$Id: analysisUtils.py,v 1.1872 2014/08/16 20:30:48 thunter Exp \$ Reduction Start Date : 2014/08/18 Reduction Completion Date: 2014/08/20 Analyst Name(s) : D. Petry Project code : 2013.1.00020.S Project title : Hydrides as diagnostic tools for the z=0.89 absorption toward PKS 1830-211 Contact Scientist: Ivan Marti-Vidal : S. Muller PΙ : PKS1830-211 b 07 TE SB Name EBs in calibrated dataset: uid A002 X86fcfa X14a6 Observation Dates and Total Integration Time (from listobs of concatenated dataset): Data records: 865920 Total integration time = 2952.1 seconds Observed from 18-Jul-2014/03:26:09.1 to 18-Jul-2014/04:15:21.2 (UTC) Number of good antennas: 29 Band(s) : 7 TDM/FDM/both : FDM

An SB Checklist example (sections A, B, C)





Comments:

D) ACA - Main Array Combination -----[]

An SB Checklist example (section D)





```
Yes No
  Needed tapering to adjust beam?
                                                    [ ] [x]
  Made mixed mode (TDM and FDM) continuum image:
                                                    [ ] [x]
    Adjusted weights using weights.py (CASA4.2.1)
                                                   [ ] [x]
  For all images (also continuum):
   Set outframe correctly: BARY (extragal objects) [x]
                           LSRK (gal. objects)
                                                    [ ]
   Used pbcor = False
                                                    [x]
   Added impbcor to produce *.pbcor.fits image
                                                    [x]
       and *.flux.fits image
                                                    [x]
  Solar system object:
                                                    [ ] [ ]
      attached ephemeris with fixplanets
      cvelled to outframe SOURCE
                                                    f 1
  Comments:
continuum: RMS 2.7 mJy/beam, beam 0.36" x 0.32", peak 1.11 Jy, dyn range 411
Line: not detected
RMS 4 mJy/beam, beam 0.38 x 0.36", peak 1.135 Jy, dyn. range 283
After one round of selfcal:
RMS 1.8 mJy (or better), beam 0.36" x 0.32" achieved in all SPWs
```

A tour of the QA2 instructions wiki



Let's go over the instructions wiki ...

http://www.eso.org/projects/alma/arc/tw/bin/viewauth/Offline/Cycle2DataReduction

An example QA2 project

Project: 2013.1.00020.S

SB: PKS1830-211_b_07_TE

EB UID: uid://A002/X86fcfa/X14a6

(see separate slides)



Your first assignment



Project: 2012.1.00978.S

SB: AzTEC-1

EB UID: uid://A002/X87c075/X4c02

http://jira.alma.cl/browse/SCOPS-770 http://jira.alma.cl/browse/SCOPS-1198

A dataset from July 2014. Manual calibration.

- get assignment email
- try to get as far as possible by tomorrow, 10:00 h, and prepare questions for meeting
- should finish by tomorrow evening

Your second assignment: pipeline-calibrated data



Project: 2013.1.01267.S

SB: XMMU_J22_a_03_TE

EB UIDs: uid://A002/X98ed3f/X627

uid://A002/X98ed3f/Xb6

http://jira.alma.cl/browse/SCOPS-1524 http://jira.alma.cl/browse/SCOPS-1650

A dataset from 5 January 2015.

- get assignment email (contains path to the pipeline product package)
- use the script "runscriptforpi.sh" to obtain calibrated MSs
- study WebLog and aot file (from project prep ticket)
- create scriptForFluxCalibration to concatenate MSs -> calibrated.ms
- create scriptForImaging.py according to requirements from aot file
- try to finish by tomorrow, 10:00 h

Optional: Your third assignment - ACA data



Project: 2013.1.00991.S

SB: NGC_1614_b_03_7M

EB UIDs: 1 EB uid://A002/X8440e0/X4b44

http://jira.alma.cl/browse/SCOPS-1084 http://jira.alma.cl/browse/SCOPS-1279

A dataset from June 2014, manual calibration.

- get assignment email
- try to get as far as possible by tomorrow, 10:00 h, and prepare questions for meeting
- should finish by tomorrow evening