Quality assurance and data delivery

Liz Humphreys and Martin Zwaan
ARC ESO Garching

EUROPEAN ARC
ALMA Regional Centre

Getting ALMA Data

- Queue based dynamic scheduling

  - Programs are composed of 30-60 min scheduling blocks

- Raw data pass through multi-tiered quality assurance

  - Combination of on-site duty astronomer, ARC staff, and automated checks

- Data proceeds to data reduction/pipeline and archiving

  - Data available from ARC (ESO) within a few days after passing QA

  - Data products (images and cubes), raw data, off-line data processing software and scripts made available to PIs

  - **Expert hands-on data reduction help from ARC nodes staff** provided on request, helpdesk also available at ESO

EUROPEAN ARC
ALMA Regional Centre

Quality Assurance

- Goal: "*deliver to the PI a reliable final data product that has reached the desired control parameters outlined in the science goals, that is calibrated to the desired accuracy and free of calibration or imaging artifacts*"

- For Cycle 1: **QA on *best effort* basis**

- Errors introduced by user-supplied parameters are outside the scope the ALMA QA

- QA Approach: Breakdown of QA into broad steps that mimic data flow.

  - Data Taking: QA0 (SB-level) & QA1 (Observatory Tasks)

  - Data Reduction: QA2 (Data Reduction / Science Pipeline)

  - Post Data Reduction: QA3 (Feedback from users)

EUROPEAN ARC
ALMA Regional Centre

Quality Assurance 0 - **QA0**

- Deals with rapidly-varying calibration/quality parameters that have to be measured within the Science SBs (or sequences of SBs)

- Done by astronomers on duty in semi-real time

  - **Atmospheric Effects:** Weather Parameters, Sky Opacity, System Temperature, Phase Fluctuations, Total Power Levels, WVR Outputs.

  - **Antenna Issues**: Antenna Gain, Relative/Offset Pointing, Focus, Antenna Tracking, Geometric Shadowing, Nutators.

  - **Front-End Issues:** RF Bandpass, Sideband Ratios, Receiver Temperatures, LO Lock Status. Connectivity Issues: Total Power levels, Delay Measurements, System Temperatures, RF Bandpass, LO Lock Status.

  - **Back-End Issues**: Total Power levels, RF Bandpass, Delay Measurements.

EUROPEAN ARC
ALMA Regional Centre

Quality Assurance 1 - **QA1**

- Deals with slowly varying parameters (>1 week) of the performance of the array that will be tackled by semi-periodic "Observatory Task" observations.

- Measured by astronomers on duty

  - **Array Calibrations**: Baseline measurements, Delays

  - **Antenna Calibrations**: All-sky pointing, Focus curves, Surface measurements, Beam patterns (including polarization observations), Relative delays between polarizations of same band

  - **Source Calibrations**: Monitoring of solar-system flux standards, and secondary quasar flux standards

EUROPEAN ARC
ALMA Regional Centre

Quality Assurance 2 - **QA2**

- QA at the level of the data reduction by the Science Pipeline or semi-interactively by the ALMA Data Reducers Team

- Science goals set by the PI can be compared with the results (i.e., SNR, dynamic range, etc)

  - **Calibration Issues**: Relative and Absolute Calibration Quality Among Datasets, Bandpass Calibration (flatness and dynamical range achieved), Overall Gain Calibration

  - **Reduction Process**: Data flagging (amount, cause, etc), Cleaning Convergence

  - **Final Data products:** SNR Achieved, Sidelobe Levels, Dynamic Range, Contamination by Bright Sources Outside FOV/aliasing, Resolution (spectral & spatial), Comparison of Deconvolution Algorithms, Residual Structures

EUROPEAN ARC
ALMA Regional Centre

Quality Assurance 3 - **QA3**

- Issues missed by all the other QA assessments

- Represent deeper problems not been detected in tests of some specific observing modes, set-ups, etc

- QA3 is all-encompassing, but most of calibration and processing problems should be picked up at QA2

- PIs will report the problems to their ARCs for evaluation

- If a problem is present, the ARCs will report the problem to DSO for resolution

- Depending on the steps required for resolution, it may require re-reduction of some of the data

Data reduction and QA2

- If the requested number of executions of an SB have been observed, the data are reduced in order to perform QA2

- Data of European projects can be reduced in Chile or in Garching

- All data are reduced using **standard procedures** for flagging, calibration and imaging

  - Cycle 0: standard CASA scripts

  - Cycle 1: calibration probably done using the data reduction pipeline

    - Science pipeline started commissioning phase

    - Pipeline integrated into CASA, the tasks will look like CASA tasks

    - Single dish pipeline and interferometric will be merged

    - Large amount of heuristics already exist

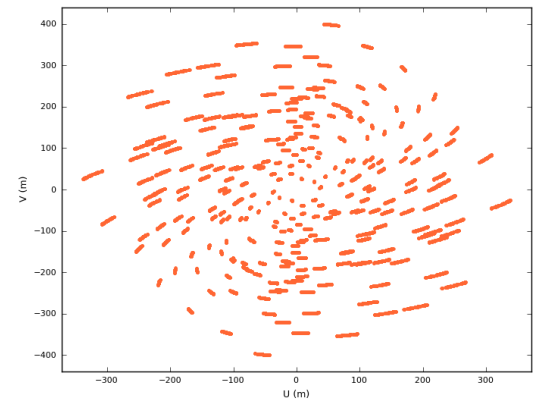EUROPEAN ARC
ALMA Regional Centre
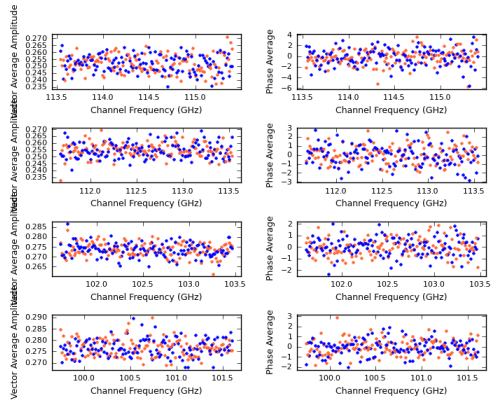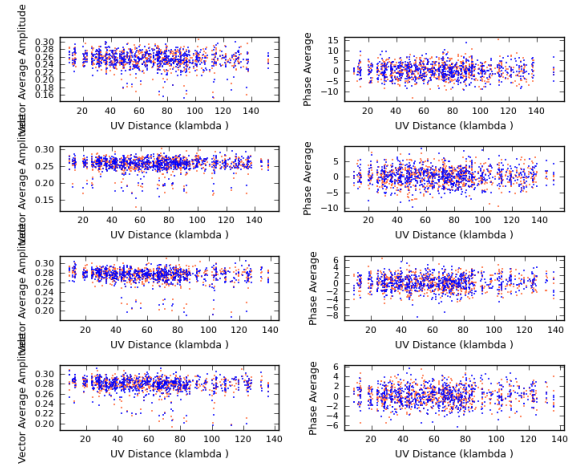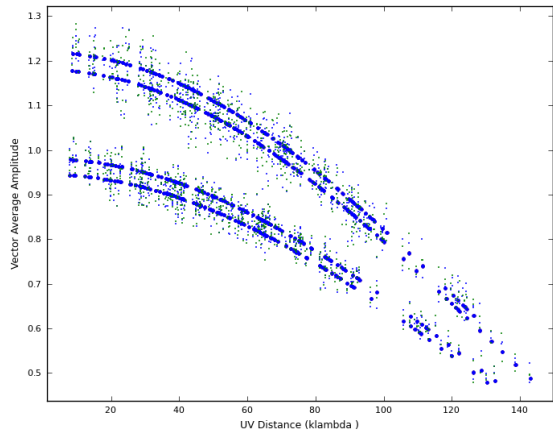
QA2 in practice

- We use *script generators* to create data reduction scripts

  - e.g., `es.generateReducScript('uid___A002_X38570d_X10.ms')`

- Some interaction/editing required:

  - antenna-based errors: bad Tsys, bad WVR, delays

  - flag atmospheric lines, bad data points

  - optimize choice of channel ranges for calibration

  - missing, resolved flux calibrators

  - etc

- In general the script generator works very well
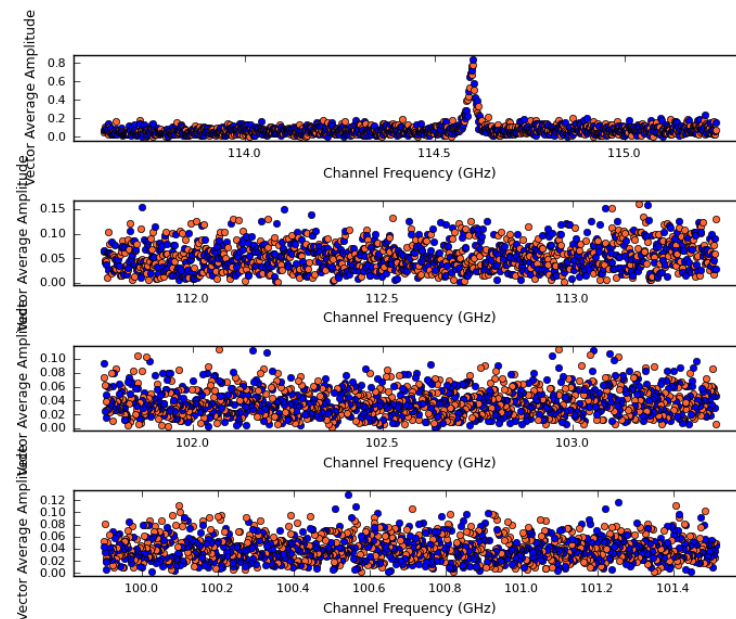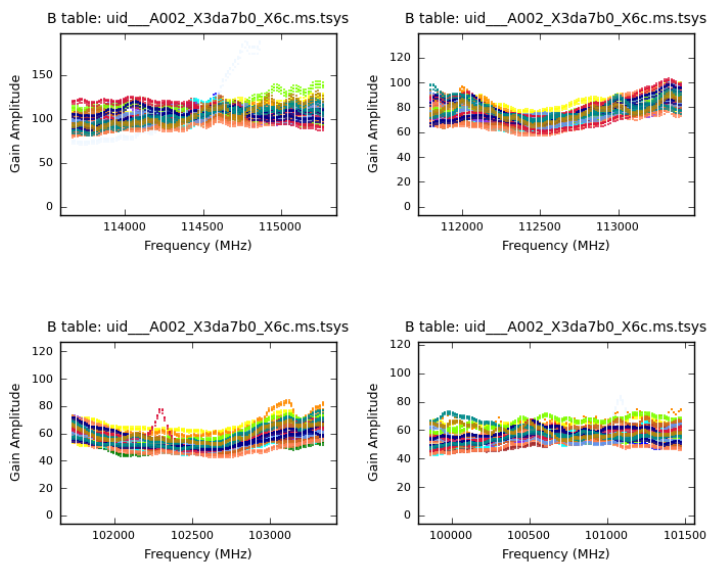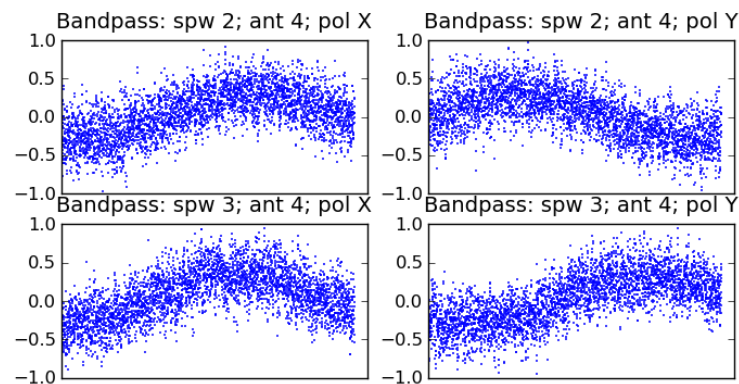
EUROPEAN ARC
ALMA Regional Centre

QA2 in practice

- After the data reduction, a number of 'qa2 scripts' are run

- These scripts are still under development

- Produce diagnostic plots and numbers related to bandpass calibration quality, WVR correction, Tsys, etc

EUROPEAN ARC
ALMA Regional Centre

Avg Bandpass for spw 0

Avg Bandpass for spw 1

Avg Bandpass for spw 2

Avg Bandpass for spw 3

Bandpass: spw 2; ant 4; pol X
Bandpass: spw 2; ant 4; pol Y
Bandpass: spw 3; ant 4; pol X
Bandpass: spw 3; ant 4; pol Y

B table: uid___A002_X3da7b0_X6c.ms.tsys
B table: uid___A002_X3da7b0_X6c.ms.tsys
B table: uid___A002_X3da7b0_X6c.ms.tsys
B table: uid___A002_X3da7b0_X6c.ms.tsys
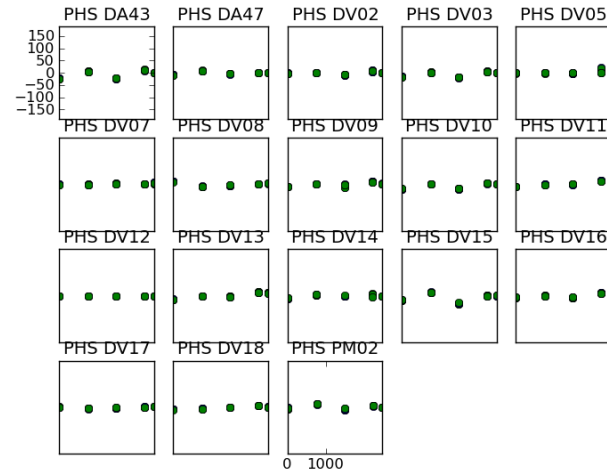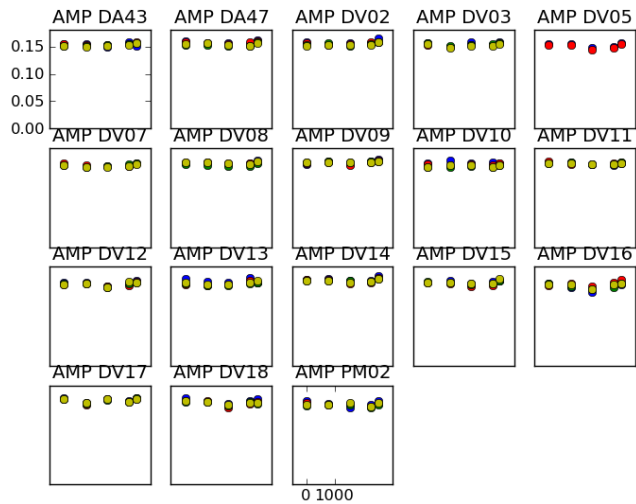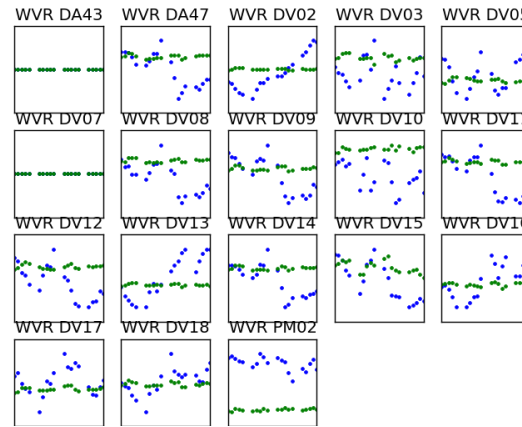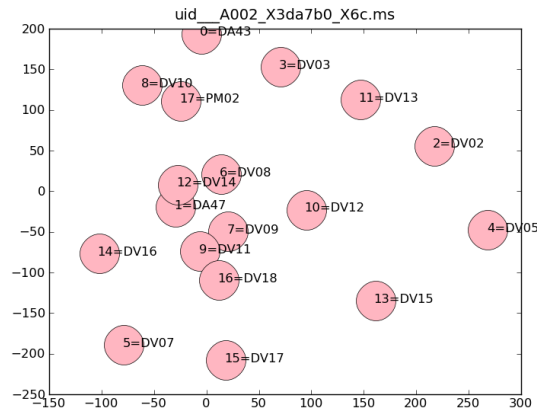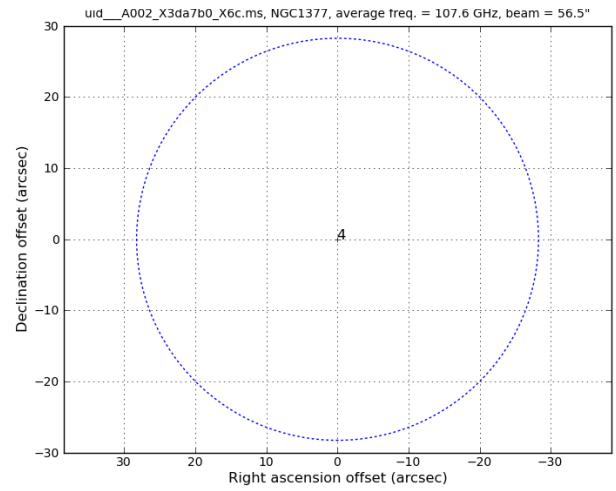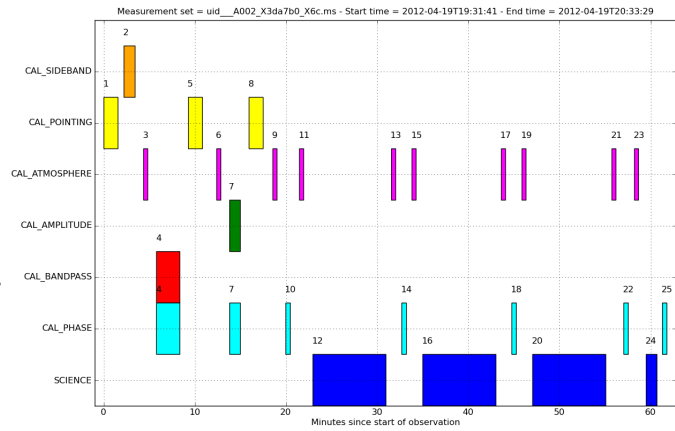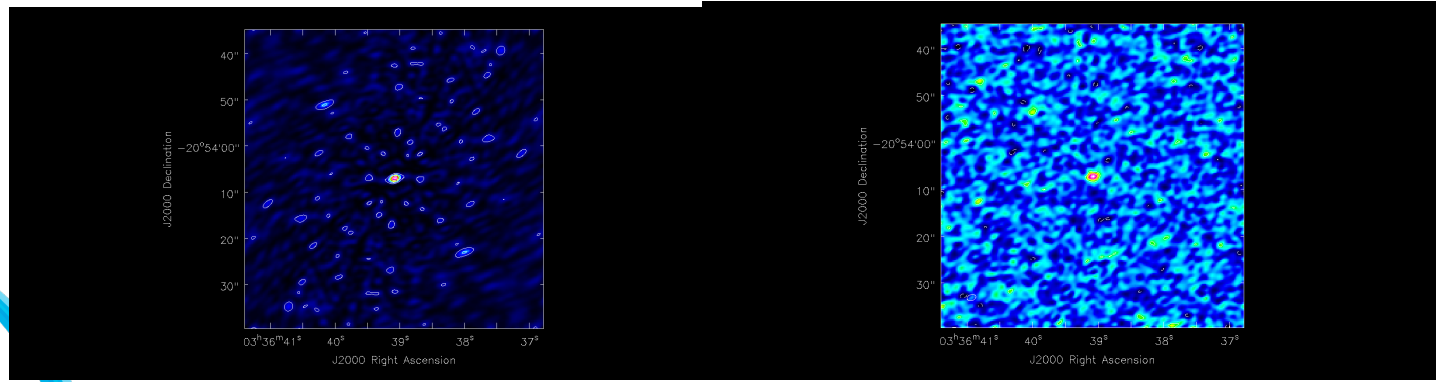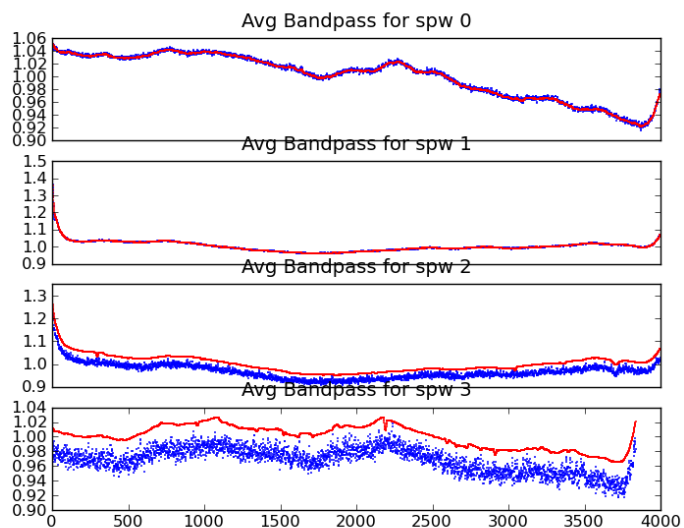
After QA2

- If data pass QA2 (rms meets requirements) data are prepared for delivery

- Otherwise, more executions of SB may be required

- Data delivery

  - Cycle 0: delivering self-consistent tar files that contain the a-priori calibrated MS, the fully calibrated MS and the/some FITS cubes. Delivered through ftp.

  - Cycle 1: delivery of data through Science Portal

  - One year proprietary time starts after data are available to PI

- Available soon: Project tracker

EUROPEAN ARC
ALMA Regional Centre

What is being delivered

- **scripts**: a priori calibration, calibration and imaging

- **raw**: data after a priori calibration

- **calibration**: all calibration tables and associated diagnostic plots.

- **calibrated**: fully calibrated (and flux equalized) data (no self-cal)

- **product**: FITS images and cubes

- **log**: *all* logs

- **qa**: checklists, diagnostic plots and printouts.

EUROPEAN ARC
ALMA Regional Centre

After data delivery

- Get help from your local ARC node and write your paper!

EUROPEAN ARC
ALMA Regional Centre

Acknowledgement text

- Acknowledgement text for ALMA publications:

  - This paper makes use of the following ALMA data: ADS/JAO.ALMA#<Project code>. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ."

EUROPEAN ARC
ALMA Regional Centre