Title: Data Reduction Workshop Lessons Learned

NAASC Memo # 109

Adam Leroy

Date: January 3, 2012

ABSTRACT

This document summarizes "lessons learned" from the first NAASC data reduction workshop, held 1-2 December 2011. It is primarily intended to inform future workshops along these lines and so focuses on suggested improvements to the existing material.

WORKSHOP SUMMARY

The NAASC presented a full 2-day program on 1-2 December, 2011 with the aim of teaching ALMA Cycle-0 investigators the basics of reducing ALMA data in CASA. The presented schedule was

Day 1 Timeslot	What	Who
8:30-9:00	Welcome and Logistics	
9:00-9:30	Overview of ALMA	Wootten
9:30-10:30	Basics of Interferometry Reduction	Schnee
10:30-11:00	COFFEE	
11:00-11:30	What's in an ALMA Data Set When You Get It	Brogan
11:30-12:30	A Crash Course in CASA	Kimball
12:30-1:30	LUNCH	
1:30-3:30	Hands on Calibration	Keohane, Scott, Crossley
3:30-4:00	Colloquium Tea	
4:00-5:00	Colloquium	
5:30-6:00	Reconvene for Q&A Related on Hands On Session	

Day 2 Timeslot	What	Who
8:30-10:00	Imaging ALMA Data	Indebetouw
10:00-10:30	COFFEE	
10:30-12:30	Hands on Imaging	Leroy, Friesen, Hale
12:30-1:30	LUNCH	
1:30-2:30	Self-Calibration	Fomalont
2:30-4:30	Hands-On Self Calibration	Leroy, Fomalont, Friesen, Hale
4:30-5:30	Final Q&A, Open Issues, Get Feedback	

Presentations are archived at the ALMA/NAASC Google sites page and can serve as the backbone of future workshops

https://sites.google.com/site/almacommunityoutreach

Registration and the conference web page were handled very nicely by D.-C. Kim and it was clear that the participants appreciated having PDFs of the presentations available during the talks. Although we had a few dropouts, most registrants attended and 17 of the 29 participants were investigators in a highest priority Cycle-0 program.

Hands-on sessions stepped participants through ALMA data reduction, imaging, and self-calibration. The CIS staff did an excellent job setting up 30 "thin client" workstations in the Edgemont Road auditorium and NAASC staff offered one-on-one support during these labs (including Braatz, Hibbard, Indebetouw, Kimball, Marcelino, and Schnee in addition to those listed in the program). The written material for the hands-on session on day one appears at the Google site. The day 2 material appears here:

http://casaguides.nrao.edu/index.php?title=WorkshopImaging

http://casaguides.nrao.edu/index.php?title=WorkshopSelfcal

(These are unlinked CASA guides intended for internal use. They are not confidential but are also not intended for general consumption.)

Nine of the 29 participants filled out our feedback survey. These gave overall positive response to all aspects of the workshop but called out some improvements that could be made. In-person feedback from participants and a staff post-mortem lead to the following suggestions that may be useful to our next workshop, which have also been noted in the presentations on the Google site.

SUGGESTED REVISIONS TO PRESENTED PROGRAM

- *The ALMA Overview Talk Could be Condensed:* This was, in retrospect, somewhat redundant to a crowd of approved ALMA users. About 15 minutes merged with the introduction would be fine. If the audience did not consist of PIs or users then this could be re-considered.

- *Suggested Reordering of Program (Data – Calibration – CASA):* Several participants and presenters felt that the program would be most sensibly reordered to begin with a presentation on "What's in an ALMA Data Set" followed by a high-level overview of "Calibration" and then a presentation on "CASA for Calibration." This means switching the second and third talks and ensuring that the "ALMA Data Set" talk starts from scratch (which it mostly does already).

- *Include More Basic, High-Level Material Sketching the Overall Flow and CASA Data Structures:* We may have failed to present enough 'forest' compared to tree. We got feedback requesting a higher level (one or two-slide) overview of the calibration flow and how it relates to CASA files and data structures. This makes sense, a high level, few-slide sketch of calibration, caltables, data columns (or whatever exists next time around) makes sense as a starting point. These slides would be very useful. It's not obvious that they exist already.

- *Consider Requesting Preparation:* The diverse background of the participants worried a number of the presenters (who did an excellent job talking to multiple audiences). In order to set a minimum level of expertise, we might consider assigning some reading on basic interferometry or requesting that participants work through a CASA guide before coming. Realistically, we cannot expect senior people to do much work before the workshop, but asking that novice users or graduate students read material (from the online course?) on the basics of interferometry seems very reasonable. There is

simply not time to cover this in a data reduction workshop, this is more of a Synthesis Imaging School thing.

SUGGESTED REVISIONS TO HANDS-ON PROGRAM

- *Tiered (Expert/Novice) Tracks Did Not Seem Necessary:* It was a good idea to try an "advanced" and a "novice" calibration hands-on but in the end this did not seem to accomplish much. A single series of steps and scripts for all users makes sense. In practice, most users were at least novices at CASA (or would not have attended) so starting from the basics seems good.

  Give Them Scripts to Start With (and Tell Them That You Did): Feedback on the first day of hands on suggested that the users would rather have a script in front of them than start from scratch. Given the short time available this seems reasonable. These did exist, but weren't advertised heavily. We suggest to document them a bit and shift the workflow to center on these prepared scripts (while keeping the very nice higher-level handout).

- *Give a Focused Script on Calibration and One on Data Inspection:* Both staff and participants thought that explaining how one inspects data and calibration tables was worth it's own (half) hands-on session. This could come before the calibration session and give people the practical tools to follow the calibration.

- *Keep the Hands-on Sessions Separate and Tell People That This is the Case:* People were nervous that by not finishing on the first day (or the morning of the second day) they would be handicapped on the second day. This was not the case, which was a good thing. We suggest to keep the hands-on sessions separate and let the participants know up front that this is the case.

- *A Half-Lecture / Half-Hands-On on the Viewer and PlotMS Might be Good:* The CDE material for 2011-2012 will include a presenter talking and demonstrating while participants work on their laptops and ask questions. For these heavily GUI, somewhat complicated components this would be a good option to explore next time (e.g., a 15-20 minute mini-lecture during the hands on part – using canned material that everyone can access).

- *Make all Hands-on Data Public:* Participants were interested in continuing their experimentation after the workshop. They asked to obtain copies of the hands-on data. This was fine for the science verification data, but the QSO data that we used had not been released or placed publicly. For future workshops we suggest that all data used in the hands-on be available,

formally or informally, to participants when they go home. This could be via USB key and heavy caveat-ing (this is not real ALMA data) would be fine.

- *Hand them a Fully Fleshed Out Directory Structure With Scripts:* Asking participants to create directories or download scripts caused a certain amount of general chaos at the beginning of each session (generally lasting about 10-15 minutes). To avoid teaching linux, it makes more sense to just make the directory structure – complete with working directories – and populate it with all scripts before hand. In order for this to happen:

- *Final Hands-on Material to CIS 3-5 Days Ahead of Time:* This allows the material to be effectively copied to all accounts and then checked.

LOGISTICS, AND INFRASTRUCTURE SUGGESTIONS

- *Target 30 Attendees for Hands-On:* 30 attendees seems like an effective cap for the hands on portion of the workshop based on the space, hardware, and in-house support available. If demand becomes very high we suggest more frequent tutorials. The Synthesis Imaging Summer School may also offer a good alternative for many users.

- *Avoid the Colloquium When Scheduling*: Some physical set up, scheduling, and hands-on difficulties arose from having to accommodate the NRAO/UVa colloquium in the same room as the workshop on the afternoon of the first day. We strongly suggest avoiding overlapping the Thursday colloquium. ER230 was suggested as an alternate venue but the infrastructure is no there.

- *Remote Participation Was Lower Than Initially Expected:* When remote participation was initially offered a number of groups responded with interest. However, despite a very nice set up led by Jonathan Keohane and David Halstead, only a few (~2) people participated remotely. We speculated that the early start time (~8:30am EST) suppressed participation by western participants. This could be addressed with a later start time, perhaps making the workshop three days (afternoon-day-morning). The issue might also be rendered moot by good "webinar" or other online material that could be accessed on the participant's schedule. If this material exists, offering remote participation should not be a priority. In any case, the consensus remains to not attempt to support remote hands-on material.

- *Physical Computer Set Up:* The consensus was that more physical space between computers and between rows was desirable. This would allow easier access for staff during the hands-on session and would allow participants to set up their laptops more easily. Josh Malone thought that

expanding the set up so that the 30 computers spanned 5 rows with lots of space would be straightforward and sensible.