ALMA Offline User Test
February 2004
Final Report and Testing Summary

ALMA-70.65.00.00-001-A-RPT
Version: A
Status: Draft
2004-02-27

Prepared By:
Name(s) and Signature(s) | Organization | Date
---|---|---
Debra Shepherd | NRAO | 2004-02-25

Approved By:
Name and Signature | Organization | Date
---|---|---
Brian Glendenning | NRAO | 

Released By:
Name and Signature | Organization | Date
---|---|---
## Change Record

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Affected Section(s)</th>
<th>Change Request #</th>
<th>Reason/Initiation/Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2004-02-25</td>
<td>all</td>
<td>ALMA-00.00.00.00-000-A-CRE</td>
<td>First Draft</td>
</tr>
</tbody>
</table>
Table of Contents

1 EXECUTIVE SUMMARY ................................................................. 4
2 TEST SCOPE .................................................................................. 4
3 TEST DEFINITION AND PREPARATION .................................... 5
4 TESTERS ....................................................................................... 7
5 TEST RESULTS .............................................................................. 7
6 AUDIT OF SCIENTIFIC REQUIREMENTS .................................... 9
7 SUMMARY .................................................................................... 10
8 APPENDIX A: A SUMMARY OF QUESTIONNAIRE RESPONSES ...... 11
9 APPENDIX B: BUG REPORTS AND ENHANCEMENT REQUESTS ...... 29
  9.1 Bugs in functionality: ............................................................... 29
  9.2 Problems with documentation: ............................................... 32
  9.3 Enhancements identified (based on Audit comments): ............ 33
10 APPENDIX C: USER TEST AUDIT .................................................. 36
1 Executive Summary

The ALMA Offline software (Astronomical Information Processing Software, AIPS++) was tested by outside users to determine if recent software improvements were adequate to meet ALMA scientific requirements. The tests focused on functionality associated with single field data. Usability of the software was not graded (e.g. Glish command line formatting and processing speed), however, note that processing speed is being addressed in a separate benchmark program.

Four datasets (2 VLA & 2 PdBI) were reduced and imaged by five testers with experience in interferometry. Three of the testers had limited to extensive experience with AIPS++. The remaining two had experience with GILDAS/CLIC, MIRIAD, and/or AIPS. All testers were able to fill, edit, reduce, image, and analyze their datasets from end-to-end in AIPS++ and obtained scientifically valid results. A total of 106 scientific requirements were exercised sufficiently to determine a grade. Of those, 57% of the requirements were deemed 'Adequate'; all necessary functionality was available and worked properly. An additional 27% obtained the grade of 'Adequate but Enhancement Desired'; the function worked properly but it would be nice to have enhanced functionality beyond the basic operations available. 12% of the requirements were graded as 'Inadequate'; the function was not possible or was severely limiting. The most significant improvement since the first audit of the Offline software is seen in the number of requirements deemed Inadequate - the number of 'I' grades decreased by more than 50%. Most notably, all requirements with 'I' grades that prevented processing of the data, were resolved.

Based on the requirements Audit for this ALMA user test, the success of each tester in reducing and imaging their data, and on the detailed answers to the questionnaire, this test is considered a success. In particular, the additions in the functionality for ALMA-like data and the significantly increased robustness of the software along the tested paths show marked improvement since the time of the IRAM/Phase II tests (ALMA Memo 473).

2 Test Scope

The primary focus of this test was to evaluate the underlying functionality of data reduction and basic imaging steps for single fields using the AIPS++ software package and determine if the functionality will likely meet ALMA scientific requirements. This includes filling the data, basic editing, calibration, deconvolution, and some basic image analysis needed to evaluate the test results. The documentation (both cookbook and online reference manual) should be adequate to figure out how to proceed with the test (this assumes some basic knowledge of radio interferometry techniques). GUIs were only used when the user had to interact with the software (e.g. identifying and flagging bad...
data, interactively defining deconvolution regions, viewing the final image and obtaining statistics).

The AIPS++ software is in an active stage of development. The current infrastructure, GUIs, and scripting language (all Glish-based) will be replaced, leading to a change in the overall user interface. A prototype framework is being built and evaluated but is not ready to test at this time. Thus, this test focused only on testing the underlying functionality of the C++ code to determine if it was adequate to meet the needs of ALMA. The test used Glish scripting to reduce and image the data. GUIs were only used when the user needed to interact with the software (e.g. identifying and flagging bad data, interactively defining deconvolution regions, viewing the final image and obtaining statistics). Development of the existing Glish-based GUIs that were not essential to the data reduction paths in this test has been discontinued while priorities changed to support the framework conversion and added functionality required for millimeter data reduction. Note: user interface elements will be tested in 2005.

Test results are important for several reasons:
1. Issues identified during the tests will help drive software priorities to develop functionality for ALMA offline data reduction and imaging.
2. This is the first outside ALMA user test which will grade the science software requirements to determine if the software meets the functional needs for ALMA.
3. Scripts generated by users will continue to be used during automated regression testing and will provide valuable insight to the pipeline heuristics team who will use the detailed processing steps and parameter inputs to develop automated pipeline decision trees.
4. The results of the testing will be evaluated in the next Critical Design Review to determine how the Offline subsystem development should proceed.

3 Test Definition and preparation

Four datasets were chosen to be processed through AIPS++: two spectral line Plateau de Bure Interferometer (PdBI) projects which observed 1 and 3mm line and continuum emission (a carbon star with a CO ring and a young stellar object with many 1 & 3 mm lines and continuum); and two 13 mm spectral line Very Large Array (VLA) projects (NH3(1,1) & (2,2) toward two star forming regions). The PdBI data represented low S/N data that may be appropriate for ALMA sub-millimeter observations while the VLA data represented high S/N data with a large number of baselines that should have similar processing steps to ALMA millimeter observations.

The two PdBI datasets were also test data for the IRAM/AIPS++ 'Phase II' comparison that was done in March 2003 (ALMA Memo 473). At that time, the Carbon star project could be reduced by the 'iramcalibrator' tool that had been developed specifically for that
test. However, the imaging was not done in the limited time due to a bug with interactive masking. The young stellar object project could also be reduced with the iramcalibrator tool however, limitations with uv-continuum subtraction, and spectral line selection prevented imaging. Before this test, the iramcalibrator tool functionality was migrated to the standard AIPS++ calibrator tool, thus, the processing for this test exercised a different reduction path than those in the Phase II comparison tests and full imaging was exercised.

Before the ALMA test began, internal NRAO scientists reduced and imaged the selected datasets. The test began when at least 2 people (one or more inside of the AIPS++ development group and one outside of the group) could fill, calibrate, and image the data from end-to-end and produce scientifically valid results. The reduction scripts and images generated during this “pre-testing” were then available to the AIPS++ group for user support (e.g. if a user had a problem their scripts could be easily compared with ones on file to find the problem quickly). This preparation minimized the probability of finding serious bugs in the software and decreased response time to users. Note, the testers were not given these pre-testing scripts to reduce their data, testers were expected to read the cookbook and on-line documentation to develop their own scripts or adapt the example scripts from the cookbook to the needs of the given test data.

A web site was developed by the AIPS++ development group to centralize information relating to the test and help testers understand their assigned data. The site, located at http://projectoffice.aips2.nrao.edu/almatst1/ALMA_TST1.html contained the following information:

1. All background information on the test scope and e-mail contacts.
2. Blank questionnaire and requirements sheets.
3. Information about how to download and install binary versions of AIPS++ for each operating system being exercised (Linux Redhat9.0, Vine, Suse).
4. A detailed cookbook and text scripts for data exercised in the cookbook. The text scripts illustrated steps in the cookbook and could be used as a starting point for users developing their own scripts for different data.
5. Dataset summaries.
6. A testing comparison page with sample images and spectra for each dataset (except one that was proprietary). The images were generated by internal NRAO testers before ALMA testing began.

Testing was expected to take 40 hours or less. Testers had 1 month to install AIPS++, complete data reduction and imaging (writing full scripts of all tasks executed), fill out the questionnaire and grade scientific requirements that were exercised. The test began on 19 January 2004 and ended on 18 February 2004. AIPS++ software developers were available to help with any questions.
The final detailed scripts generated during this test will be given to the ALMA pipeline heuristics team. Thus, the step-by-step processing decisions of expert interferometrists participating in this test will help define the heuristics for automated pipeline processing of ALMA data.

4 Testers

The group of testers were chosen to have interferometric data reduction experience in at least one of the following reduction packages: AIPS, AIPS++, MIRIAD, GILDAS/CLIC with data from the VLA, BIMA, and/or PdBI. Testers included those internal to the ALMA project as well as external individuals:


Chris Wilson - McMaster University, SSR Subsystem Scientist for the Pipeline Subsystem, member of the Science IPT, and participant in the IRAM/AIPS++ Phase II tests in 2003.

Munetake Momose - Ibaraki University, Observer of the ALMA Science Advisory Committee (ASAC).

Crystal Brogan - University of Hawaii, External tester, former member of the NRAO science testing group for AIPS++.

Friedrich Wyrowski - Max Planck Institute, External tester

5 Test Results

All testers were able to fill, edit, calibrate, and image their data. Testers used simple image analysis functions to evaluate the scientific quality of their data - all testers obtained scientifically valid results.

All testers reported having adequate information and documentation necessary to complete their testing. The primary improvement suggested was a better User Reference Manual written more for the end-user. Testing support was considered very helpful to excellent.

Testers spent from 17 to 38 hours on the testing process (from AIPS++ installation to final evaluation of the requirements). The actual calibration, imaging, and analysis steps took between about 15 hrs and 31 hours. The longer processing times were generally from users who had little or no prior knowledge of AIPS++ (and so had to learn the
software from scratch) and/or did significant imaging and analysis tests (e.g. mosaiced different fields, did primary beam correction, made moment maps and took spectra at different locations).

The test scope was limited to evaluating the functionality of the basic C++ code. Thus testers did not, in general, submit problems associated with the GUIs or Glish tasking code. A number of functional problems were identified during testing - none were severe enough to prevent end-to-end processing. A detailed list of all bugs found during the test is included in Appendix B. The most serious problems were that splitting out the data and doing continuum subtraction in the uv plane did not work with the Vine version of AIPS++ (a Linux operating system used in Japan but not readily available in the US). For the split function, the AIPS++ group found the problem, re-built the software, and made the new RPM available to the tester within a few days (the fix allowed the data to be split but was still 'quirky' and needs additional work). The problem associated with uv continuum subtraction was found at the end of the test. The tester informed the AIPS++ group of the problem after the test, and subtracted the continuum in the image plane. All problems encountered were minor and should be relatively easy to fix during the next development cycle. The Subsystem Scientist and AIPS++ development Lead will work out a schedule to correct problems before the next ALMA user test.

The scripts generated during this exercise will be used as regression tests to ensure that problems along the exercised paths through the software are fixed and no new ones are introduced.

Despite the intensive nature of this test, the overall grade of the testing experience ranged from fair to excellent. One external tester provided this insight into the testing experience:

- It is very difficult to compare this AIPS++ testing experience with other data reduction packages. For every package one needs time to get to know it and all its functionalities. During the test, obviously, I got increasingly faster in using AIPS++ and also in finding out about the available functionalities. I remember that my first encounters with classical AIPS were full of frustration. This AIPS++ test was certainly not that bad and I would already now prefer AIPS++ over AIPS in terms of using the package and also in some of its functionality. On the other hand, AIPS++ has still to improve its speed, functionality, simplicity of usage quite a bit to replace MIRIAD as my mostly used interferometric data reduction tool. I want to stress that while there are performance problems and still lots of areas for improvements, for VLA data I prefer already now using AIPS++ instead of classical AIPS. In the current state, AIPS++ seems to be better adapter to handle VLA data than millimeter interferometer data (e.g. my PdB data set).
6 Audit of Scientific Requirements

Testers were given a list of the ALMA Priority 1 requirements for the Offline Subsystem that could have been exercised (all or in part) by this test. They were asked to evaluate each requirement based on the following scale:

- A Adequate
- A/E Adequate but some enhancement is desired
- I Inadequate
- U Unable to grade (e.g. function not exercised or data did not require this step).

If grades A/E or I were given testers were asked to provide a brief comment on why the requirement was not adequate. Appendix C provides a summary of the testing grades along with a combined grade for the test. Appendix B also lists all enhancements that were mentioned for requirements given A/E grades.

110 Priority 1 requirements could be graded based on functionality only (out of 273 possible requirements). Of these, sufficient responses were obtained on 106 to provide a reasonable estimate of the grade for this test. 57% of the requirements were deemed Adequate; all necessary functionality was available and worked properly. An additional 27% obtained the grade of ‘A/E’; the function worked properly but it would be nice to have enhanced functionality beyond the basic operations needed. 12% of the requirements were graded as Inadequate; the function was not possible or was severely limiting. None of the requirements graded as Inadequate prevented processing of any dataset.

In comparison, the first ALMA Offline Subsystem Audit based mostly on documentation with some limited testing of VLA data showed that 66% of the requirements received grades of ‘A’, 6% had ‘A/E’ grades, and 26% had ‘I’ grades (percentages based on the 110 requirements that overlap with this test evaluation). The table below summarizes the comparison for the subset of requirements that were tested:

<table>
<thead>
<tr>
<th>Grade</th>
<th>First Audit Results</th>
<th>TST1 Audit Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>66%</td>
<td>57%</td>
</tr>
<tr>
<td>A/E</td>
<td>6%</td>
<td>27%</td>
</tr>
<tr>
<td>I</td>
<td>26%</td>
<td>12%</td>
</tr>
<tr>
<td>U</td>
<td>2%</td>
<td>4%</td>
</tr>
</tbody>
</table>

As expected, once the requirements were tested with actual data (both VLA and PdBI data), the number of requirements with A/E grades increased, e.g. testers found some function lacking or some addition would be nice. The percent of requirements with A or A/E grades increased from 72% at the time of the first Audit to 84% during TST1.
Note that PdBI data could not be processed in AIPS++ when the first Audit was made. Adding a new calibration path to reduce low S/N millimeter data (including phase transfer from 3 mm to 1 mm) significantly increased the functionality of AIPS++ while increasing the number of requirements with grades of A or A/E.

The number of requirements that could not be graded remained at about the same level (e.g. a few each time). The most significant improvement is seen in the number of requirements deemed Inadequate - these decreased by more than 50%. Most notably, all requirements with I grades that prevented processing of the data were resolved (e.g. User cookbook available, spline gain solutions and polynomial bandpass solutions for low S/N data, 3 mm to 1 mm phase transfer for PdBI data, antenna gain and atmospheric opacity corrections, uv continuum subtraction, blinking between images).

A list of all 273 Priority 1 scientific requirements (with shortened description) is provided in Appendix C. To the right of each requirement are the 1st Audit grades for comparison, the final grades from this test for 110 requirements that were exercised, all individual grades, and comments from testers.

7 Summary

Based on the Audit of the requirements in this ALMA user test, the success of each tester in reducing and imaging their data, and on the detailed answers to the questionnaire, this test is considered a success by the ALMA Offline Subsystem Scientist. In particular, the additions to the functionality for ALMA-like data and the significantly increased robustness of the software along the tested paths show marked improvement since the time of the IRAM/Phase II tests (ALMA Memo 473).
8 APPENDIX A: A Summary of Questionnaire Responses

Responses are coded by the tester’s initials and color:

CB = Crystal Brogan, MM = Munetake Momose, DS = Debra Shepherd,
CW = Chris Wilson, FW = Friedrich Wyrowski.

1. Please list briefly your background in the following areas:

A. Radio Interferometry ((sub)millimeter or centimeter)

CB: I have used the VLA extensively as well as the VLBA, ATCA, GMRT,
and BIMA.

MM: I have long experience in reducing data taken with the Nobeyama
Millimeter Array (NMA). I have some experience in installing a
new correlator and developing reduction software for the NMA.

DS: I've reduced OVRO 1 & 3mm data, VLA mm & cm data, PdBI
millimeter data. I have worked extensively with OVRO data,
mostly CO or SiO mosaics & continuum. I've also worked with
BIMA data (combining 12m+BIMA single field).

CW: I have used the Caltech Millimeter Array extensively and also
observed with the JCMT-CSO single baseline submillimeter
interferometer. I have a little experience with BIMA and the
VLA

FW: Observing planning, data reduction and analysis from 1993 on
BIMA Postdoc 1998-2001

B. Experience with VLA and/or PdBI data

CB: I have reduced numerous VLA data sets ranging from 74 MHz to 23 GHz.

MM: No experience.

DS: I've been reducing VLA data for about 6 years (at first using
AIPS, now using AIPS++ exclusively). I have reduced 3 PdBI
datasets in the past year, all in the AIPS++ package.
**CW:** My only experience with PdBI data is from the IRAM-AIPS++ Phase II test in February 2003

**FW:** VLA: 8-45 GHz, B-D arrays, spectral line + continuum
PdB: about 5 PI projects, many times CoI with involvement in data reduction

**C. Astronomical Data Reduction packages:**

**CB:**
- AIPS       Extensive
- MIRIAD  Moderate
- AIPS++   Moderate

**MM:** Experienced user of AIPS, but no experience for the other packages. I have also experience in the installation of AIPS.

**DS:**
- AIPS      Used occasionally for the past 6 years
- MIRIAD    Used extensively for imaging OVRO data
- MMA       Used to reduced OVRO data
- Gildas/Clic Attended a Clic tutorial in early 2003, have never reduced data in Gildas/Clic
- AIPS++    Testing AIPS++ for about 3.5 years, for testing preparations, I reduced and imaged the other datasets (2 VLA and one other PdBI) but did not look at the dataset I was assigned.

**CW:** I'm very familiar with the imaging parts of MIRIAD and with MMA. I've used AIPS but am very rusty. I've never used Gildas/Clic. I used AIPS++ for the Phase II tests.

**FW:**
- AIPS       yes, for VLA data
- MIRIAD     yes, for BIMA/VLA data
- MMA        what is MMA?
- Gildas/Clic yes for PdBI data
- AIPS++     a year ago I reduced a VLA continuum observation out of curiosity with AIPS++. Worked.

**D. How much experience have you had with the AIPS++ software package before this test?**

**CB:** I was a member of the NRAO Aips++ user group committee for two
years. As part of that group I learned the basics of Aips++, tested some parts of it -- focusing on visualization, helped to write/comment on the ALMA external software requirements document, and was a member of the team that carried out the first AIPS++ ALMA audit.

**MM**: No experience.

**DS**: Significant. I’ve been reducing all my VLA science data with AIPS++ for the past 2 years.

**CW**: I spent about 40 hour working with AIPS++ during the IRAM-AIPS++ Phase II test, including learning the package and the IRAM-specific functions from scratch, in February 2003.

**FW**: A year ago I reduced a VLA continuum observation with AIPS++.

2. **Please identify which dataset you processed during this test:**

- **CB**: VLA NH3(1,1) observations of G192.16
- **MM**: VLA NH3(1,1) & (2,2) observations of NGC 7538
- **DS**: PdBI 1 & 3mm line observations of IRAS 20126
- **CW**: PdBI 1 & 3mm line observations of U Cam
- **FW**: PdBI 1 & 3mm line observations of U Cam and VLA NH3(1,1) & (2,2) observations of NGC 7538

3. **Were you able to complete the fill, editing, and calibration of the data? If not, why? Please comment on specific steps if desired (comments can be positive or negative):**

   **A. Filling the data into AIPS++ format**

   - **CB**: Yes. No problems after it was realized that async=F would make the filler finish loading one file before beginning the next.
   - **MM**: Yes
   - **DS**: Yes
CW: I was able to complete all steps. I was only able to process the first day's worth of data as I ran out of time due to other commitments during the testing period. The data I worked with were 05may98

FW: PdBI: yes, no problem.
   VLA: I first forgot to correct the DATA_DESC_ID as described in the cookbook, which led to some confusion. As said in the cookbook, a future version of the vlafiller should take care of this.

B. Editing and visualizing your data

CB: I found this a bit difficult due to the large size of the G192 dataset. Each new "view" of the data took several minutes to load and was thus very time consuming. The current inability to flag on an averaged "CH 0" dataset (especially for the calibrators) or to flag one channel and have it applied to all channels also contributed to prolonging the editing process in a painful way. Use of msplot (to identify) and flagger (to flag all channels) was a doable and fairly easy, if rather time consuming, compromise.

MM: Yes

DS: Yes

CW: Yes

FW: PdBI: Editing was ok. Visualizing could have been better. I'm used to look also at time averaged uv data during the various calibration steps which is important for low S/N data. This was not possible with msplot. E.g. looking at the average of all uv data in a given correlator window tells you usually immediately whether there is a line or not (if situated roughly toward the phase center).

Another issue with visualization is that all the tasks are bound to a measurement set on which they operate. But sometimes one wants to switch back and forth between different observing days or calibrators, which might have different ms files and then one has to open another msplot or viewer and it seems that the save&restore of the parameters used for a given
view at the data doesn't work correctly. And opening just several msplit/viewer for the different ms files will swamp you with windows...

Where can I see the System temperatures, optical depths etc. of the observations? This would have been helpful to compare different days. Probably the weights are proportional to $1/T_{\text{sys}}^2$ but still just to read somewhere whether $T_{\text{sys}}$ is 200 or 1000K gives a feeling of what the weather was like.

VLA: Interactive editing worked fine. I had a very brief go at using the autflag routines but couldn't get it to work. For VLA data clever visualization is even more crucial than for PdB due to the many more baselines. There should definitely more averaging options for the display.

C. Atmospheric phase corrections (PdBI data only)

**CB:** N/A

**MM:** N/A

**DS:** Yes

**CW:** Yes

**FW:** PdBI: There wasn't much to do. It was just a parameter given during the filling and later on there is no control over applying the correction or not. I actually loaded one data set out of curiosity with DOPHCOR=F and couldn't see any difference at all to the data with DOPHCOR=T. Switching between applying the correction or not, for any part of the data set, should be made very easy.

VLA: N/A

D. Gain (phase and amplitude) calibration

**CB:** No problems. The setup is a little convoluted, but it was easy to change the cookbook example to be applicable to g192.

**MM:** Yes

**DS:** Yes
CW: Yes

FW: PdBI: I would have liked pgplot plots instead of the ps files which have to be looked at with ghostview. Also, the phase noise for a given antenna solution should be given into the log window in degrees. I know the fit results can be found in the log files but this information is so important that it should go into the log window.

Controlling the splines is not as easy as using polynomials. For very noisy data, I would have liked to use low order polynomials so that the fit doesn't simply follow the noise. Probably there should be both methods to do fits. I really like the MIRIAD GFIDDLE method where you can change orders of polynomials interactively, while flagging bad data at the same time. And this can even run in a batch mode (All thanks to Peter Teuben).

For the 1mm data, I tried to apply the scaled raw 3mm phases instead of the splines but couldn't get it to work with the information on the web (Although the cookbook writes this is the recommended method it only describes using the 3mm splines for this correction).

VLA: Worked fine, but I would like to see phase (degrees) and amplitude rms for each antenna in the default logger.

E. Absolute flux calibration

CB: The flux calibration itself proceeded with no problem, however, it is at this stage that the "MODEL" data column get created which takes >10 minutes for this large dataset.

MM: Yes

DS: Yes

CW: Yes

FW: PdBI: Absolute flux calibration with noisy data as the one I had is always a pain. There are different philosophies of how to handle decorrelations and there should be more control and flexibility of how to derive the fluxes of the secondary calibrators. The decorrelation has to be taken into account.
To me it seemed some information on the project summary sheet where wrong, e.g. the given fluxes of 3C454.3. Since for day 2 the MWC349 measurement was useless, I used fluxes from day 1 to fluxcal day 2.

Is it possible to unset fluxes set in the data files? This would be helpful for trying out different flux cal variants. I ended up having backed up the "raw" measurement sets and then copying them back to start a fresh flux cal.

VLA: Worked fine.

**F. Bandpass calibration**

**CB**: No problems

**MM**: Yes

**DS**: Yes

**CW**: Yes

**FW**: PdBI: No problem

VLA: No problem

4. *Were you able to subtract continuum in the uv-plane (if applicable)? If not, why?*

**CB**: Unnecessary.

**MM**: NO, I've tried, but was not able to do. I have no time to ask the developers about this issue, because the end of testing campaign was so close. I made continuum subtraction in image plane.

**DS**: Yes, I did uv continuum subtraction for the 1mm data (it was not needed at 3mm). The process worked well.

**CW**: N/A

**FW**: PdBI: Was not needed for my data set.

VLA: The continuum in this data set was very offset from the phase center at the edge of the primary beam. For a proper uv
subtraction in such a case it is needed to shift the data to the center of the continuum emission for the subtraction (see MIRIADs UVLIN, offset option). This is not possible. Therefore I continued without the continuum subtraction. Since the bulk of the line emission is offset from the continuum anyway, this was not a problem, but rather a nice check of the reduction of channels maps in which the compact continuum then always showed up.

5. **Were you able to split out the calibrated data (if desired)? If not, why?**

**CB:** The new splitting task worked fine.

**MM:** Yes, but the behavior of the "split" function was tricky (even in the updated version). Please check out my detailed report and scripts.

**DS:** Yes, no problems.

**CW:** Yes

**FW:** PdBI: No problem.

**VLA:** No problem

6. **Were you able to image the data? If not, why? Please identify any problems you had during imaging.**

**CB:** -

**MM:** Yes, I got several images.

**DS:** Yes, the 3mm imaging was relatively easy. The 1mm continuum bands had more than 10 strong lines contaminating the bands, these had to be laboriously identified in the combined image cubes and then subtracted before obtaining a continuum only image. This was not particularly easy but it would not have been easy in any software package I know of.

**CW:** Yes, although I did not experiment at all with clean boxes because of being short of time. I had a little difficulty at first being sure that the images were being produced correctly. However, this was because I had not imaged enough channels and so the emission that I saw didn't look like what I expected (pilot error).
FW: PdBI: Yes. For my source it would have been helpful to define for each channel a different spherical mask but I couldn't figure out how to do it. Instead I used imager.maskregion where I specified a polygon on a preliminary image. The documentation was not very clear about how to do it and there are lots and lots of details about different masks/regions where it was easy to get lost. Since I used MEM deconvolution, I couldn't just interactively specify masks.

VLA: Yes. One thing I didn't manage to do is to average the continuum from both IFs to increase the S/N in the continuum map. I tried something along the lines of

```python
imgr.setdata(mode='channel', nchan=[1,1], start=[1,1],
             step=[1,1], spwid=[1,2])
imgr.setimage(mode='channel', nchan=1, start=1,
              step=2, spwid=[1,2]);
```

but it didn't work (see in my macros DO_SPLIT2 DO_SMAP2) error message: gridding routines could not find any overlap in frequency between image and visibility channels in spw 2 check the data selection and image definition

Since the continuum was at the edge of the primary beam, another restriction was that I couldn't specify for the primary beam correction using setvp() the radius to which to apply the correction.

For the processing of the Mosaic I used mosaicwizard and the multi-scale clean algorithm which seems to be a very nice tool to recover in large fields also the large scale emission. On the other hand, in the data flux was clearly missing (see the discussion in Zheng et al 2001, they added Haystack data and then obtained much better results), leading always to negative bowls in the deconvolved images.

A desired feature: an option to overwrite existing clean models

7. Were you able to analyze the images adequately to determine if the results you obtained were scientifically reasonable (e.g. display the image, calculate RMS and peak, make a moment map or take a spectrum)? If not, why?
CB: I didn't have any problems, but I should add that this is my main previous area of expertise with aips++. ie. I already knew how to do it.

MM: Yes.

DS: Yes, I did minimal analyzing to check the scientific correctness of the results. I took spectra, obtained RMS and peaks, calculated total emission in the regions. I did not make a moment map - no need, the emission was compact.

CW: Yes. I found the viewer a bit hard to learn, but for the purposes of this test I was able to do what I needed to do.

FW: PdBI: Yes, the image analysis tool which can be called from the default viewer is helpful. I couldn't determine spectra integrated over a given polygon. All the options to do it are there, but either the double click into the polygon is not accepted or the display for the integrated flux is not properly scaled. An option for giving manual plot limits would be good.

I would have liked to change the reference frequency of the data cubes to the CO 1-0 and 2-1 rest frequencies so that the velocity axis is Vlsr of the CO lines but I couldn't figure out how to do it.

To look at data cubes using the canvas manager is somewhat complicated. I would have hoped for something more automatic. E.g. in GILDAS or MIRIAD, if I look at a cube I will see per default already more than one channel. Also for many channels one always has to change lots of parameters, as I said, saving them didn't work properly for me. I also couldn't get a velocity or channel marker at or in each box.

VLA: Yes. But see also the comments on the PdB data set.

8. Please summarize the final results of your image(s):

CB: - RMS: ~2 mJy
   - Peak Flux Density: 23.8 mJy/beam (beam 3.8" x 3.2")
     Image made with 0.5" cells; 512 pixels; cleaned inner quarter to 6 mJy.

MM: - N7538D, SPWID=3, Continuum
   - RMS: 2.5 mJy/beam
- Peak Flux Density: 94.12 mJy/beam (Beam=7."61*7."46 PA=72.8)
  - N7538D, SPWID=3, Line
    - RMS: 6.3 mJy/beam (in a typical channel)
    - Peak Flux Density: 134.8 mJy/beam at 28ch in the original
      uv data (Beam=7."61*7."46 PA=72.8)
  - N7538D, SPWID=4, Continuum
    - RMS: 2.74 mJy/beam
    - Peak Flux Density: 81.3 mJy/beam (Beam=7."66*7."53 PA=44.1)
  - N7538D, SPWID=4, Line
    - RMS: 7.2 mJy/beam (in a typical channel)
    - Peak Flux Density: 91.8 mJy/beam at 28ch in the original
      uv data (Beam=7."66*7."53 PA=44.1)

**DS:** 3mm continuum:
- RMS: = 0.47 mJy/beam
- Peak Flux Density: = 27.5 mJy/beam at 20 14 26.04 +41 13 32.52
- Total Flux Density = 38.1 mJy

1mm continuum:
- RMS: = 19 mJy/beam
- Peak Flux Density: = 198 mJy/beam at 20 14 26.04 +41 13 32.52
  peak is supposed to be at 20 14 26.0 +41 13 32.7
- Total Flux Density = 283 mJy

3mm c34s(2-1) line:
- RMS: = 12 mJy/beam
- Peak Flux Density: = 200 mJy/beam

3mm ch3oh line:
- RMS: = 18 mJy/beam
- Peak Flux Density: = 453 mJy/beam

1mm c34s(5-4):
- RMS: = 25 mJy/beam
- Peak Flux Density: = 293 mJy/beam

**CW:** I was told there was no continuum emission in this data set, so
I only imaged the lines.
- CO 1-0: peak = 0.269 Jy/beam
  true rms is 0.014 Jy/beam in line-free channels
- CO 2-1: peak= 1.558 Jy/beam
  true rms is 0.042 Jy/beam in line-free channels

**FW:** PdB1: The expanding circumstellar shell could be imaged in channels
of 5 km/s width for CO 1-0 and 2-1, although, with only two
days with lots of decorrelation at 1mm, not with the high
sensitivity and angular resolution presented in the Lindqvist et al (1999) paper. The structure of the envelope in the images resembles by and large the published results. A detailed comparison is difficult without changing the frequency of the cube to the CO rest frequencies (see note above). The spectrum toward the center position shows much higher high velocity peaks than the published spectra which probably is a resolution effect, e.g. in the paper the ratio of inner two outer shell components also changes by a factor two going from CO 1-0 to 2-1. The total 2-1 flux of order 5 Jy agree with the values from the paper. The RMS in the 5km/s line free channel is about 15 and 25 mJy/beam for 3 and 1.3mm, respectively, for channels with line emission the RMS is higher due to remaining phase errors.

VLA: The continuum emission from NGC7538 IRS 1&2 could be imaged. The flux densities of the sources of 0.27 and 1.0 Jy are lower than the values given by Henkel et al. 1984 (the primary beam correction was applied) but fluxes at the edges of the primary beam are not very reliable anyway (see ngc7538d.cont.ps, ngc7538d.cont.pbcor.ps). To image the lines it was necessary to lower the angular resolution, hence with that increasing the brightness temperature sensitivity, with appropriate taper. Both, the (1,1) and (2,2) lines are detected in the NGC7538D field (ngc7538d-spec11.ps,ngc7538d-spec22.ps toward the peak of emission) and parts of the (1,1) cube can be seen in ngc7538d.cube11.ps,ngc7538d.cube.hr.ps with different angular resolutions and for one channel as mosaic (ngc7538.line11.mosaic.ps) using the mosaic wizard.

9. Did you have adequate support during your test? If you contacted the AIPS++ groups for questions or to fix a bug, please comment on the interaction and whether it was helpful.

CB: The help I received during my test was fantastic. Each email for help that I sent was responded to very quickly and usually by multiple members of the development team (including phone calls from the subsystem scientist).

MM: The support was adequate and responsive enough.

DS: Yes, I contacted the AIPS++ group twice with problems, they responded within the hour with help.
CW: The support was very good. I always got a reply (and often the final answer) within 20-40 minutes (usually 20 minutes).

FW: I didn't ask for much support since my approach was to figure out as much as possible on my own. For two problems I contacted the AIPS++ group and got immediate and helpful responses.

10. Was AIPS++ easy to install? If not, why?

CB: Yes, it was very easy to install from the rpm's. I had a little trouble initially because even though I had not previously installed aips++, evidently I need to follow the directions for that case (something to do with pre-existing libraries). But within a few minutes of sending an email for help. Joe McMullin called me and helped me figure out the problem.

MM: Yes. It might be helpful to explicitly specified the order of installation (which RPM package should be installed first, and so on).

DS: I did not install AIPS++ - I used the stable release version available at NRAO.

CW: Very; I just had to run aupdate to update my existing version from about April 2003.

FW: Yes, with the RPMs it was very easy. The first distributed code was very large since it was compiled with debugging enabled. One problem after installation was, that it took very long to start up on certain machines. With the help of the aips++ group we found out that it was due to our /etc/hosts file.

11. Was the documentation adequate for you to complete your test?

- Was the cookbook good?

CB: The cookbook was very useful as long as one followed the prescribed data path. In order to deviate one had to go into the other documentation which could be time consuming but the information could be found eventually.
MM: Cookbook showed good examples and was very helpful.

DS: I will not answer this since I helped to write it.

CW: The cookbook was very good. Between the cookbook and the sample Glisch script that were provided, I never had to look at the on-line documentation.

FW: yes

- Was the on-line documentation good:
  * User Reference Manual?
  * Supporting documentation?

CB: The User Reference Manual (URM) has improved since the last time I tried to use it. i.e. it used to be written strictly for programmers -- i.e. you could find out the list of inputs and whether they were integers or strings but that was about it. Now at least for some functions, (I didn't check many) imgr.clean for example now gives a detailed "astronomer's description" of the different kinds of clean. Hence, you could easily decide that "Clark" clean is an option and how it differs from the cookbook example of "Hogbom" clean. Some of the inputs are still quite obscure though and need similar "astronomer's" descriptions. The examples given in the URM have also become less trivial.

Eventually, I think the aips++ cookbook will need to go to a format that is more like the AIPS cookbook where every possible input is defined AND specific data reduction paths are given as examples. The URM is still no where near as helpful as the AIPS "explain files" and I think this level of explanation is what the URM should strive for.

I also found the aips++ "VLA cookbook" had some useful examples/caveats not found in the "ALMA cookbook".

MM: On-line documentation, especially the reference manual, was also useful.

DS: Reasonable for the things I needed to know. If I did not know as much about AIPS++ I think it would be difficult to figure
out what to do.

**CW:** -

**FW:** In principle yes, but even with a search function it was sometimes difficult to navigate through lots of documentation. And in some cases the documentation referred to older aips++ versions.

### 12. Roughly how much time did you take to perform the following steps:

**CB:** - Installing aips++: 30 m  
- Fill, editing, & calibration: 10 hrs  
- Continuum imaging: N/A  
- Spectral line imaging: 8 hrs  
- Analysis: 1 hr  
- Filling out this questionnaire: 1 hr  
- Evaluating and grading the scientific requirements:  
- Total time: 21 hrs

**MM:** - Installing aips++: 1 hour  
- Fill, editing, & calibration: 24 hours  
- Continuum imaging: Only image-base continuum subtraction was successful, so this item is set to be 0 hour.  
- Spectral line imaging: 4 hours  
- Analysis: 3 hours  
- Filling out this questionnaire: 3 hours (including writing the report)  
- Evaluating and grading the scientific requirements: 3 hours  
- Total time: 38 hours

**DS:** - Installing aips++: N/A  
- Fill, editing, & calibration: 14 hrs*  
- Continuum imaging: 2 hrs  
- Spectral line imaging: 2 hrs  
- Analysis: 1 hr  
- Filling out this questionnaire: 1 hr  
- Evaluating and grading the scientific requirements: 1 hr  
- Total time: 21 hrs

* 14 hrs of reduction includes 5 hours of just trying to figure out the PdBI correlator settings (I hadn't see this data before and it was quite complicated) and 5 hours of testing
different ways of reducing the 3mm data.

**CW:** - Installing aips++: 1 hour  
- Fill, editing, & calibration: 9 hours  
- Continuum imaging: N/A  
- Spectral line imaging: 3.5 hours  
- Analysis: 2 hours  
- Filling out this questionnaire: 1 hour (including bug list)  
- Evaluating and grading the scientific requirements: 1 hour  
- Total time: 17.5 hours

**FW:** PdBI: - Installing aips++: 1 hour + several hours  
   to figure out the /etc/hosts problem  
- Total time: about 4 full days

**VLA:** - Total time: about 4 full days

13. Please rate your overall testing experience:

**CB:** good  
Overall I would say that my experience with the package itself was only fair, but the dedication of the developers in responding and finding solutions to the problems I had definitely made a big difference.

**MM:** good

**DS:** good

**CW:** excellent

**FW:** fair

14. Was the test well designed and executed by those in the ALMA offline subsystem (e.g. the subsystem scientist and the Offline subsystem group). If not, can you provide any suggestions for improving the next test?

**CB:** I thought the test was well designed, and as mentioned before the help offered was excellent.

**MM:** I think the test was well designed.

**DS:** I will not answer this question since I was one of the people
who helped to prepare for it.

**CW**: My only suggestion was that I found grading the requirements a little difficult. In some cases I didn't understand what the requirement was from the requirement description. Maybe that just indicates that I didn't exercise the requirement myself; you can judge from my evaluation form.

Also, I wasn't sure whether requirements were supposed to be graded on a scale of 1-5 or a grade of A/E/I/U or both.

**FW**: Yes. I would have liked to get the data also in PdB format for also processing it with CLIC/GRAPHIC for comparison with AIPS++

**15. Do you have any additional comments that may help improve test of the offline software in the future?**

**CB**: It would have speeded things up a bit to have had links on the http://projectoffice.aips2.nrao.edu/almatst1/ALMA_TST1.html page to the other useful pieces of documentation, like the URM.

**MM**: I feel good sample scripts were crucial for first-time users; I could not produce any image if there was no prototype.

**DS**: No.

**CW**: There seemed to be quite a number of the Test 1 requirements that I didn't exercise and so couldn't evaluate. That could be because of the data set I used OR because I was short of time and so could only do the basics. But if it turns out that MANY testers couldn't evaluate the requirements, and you really wanted the requirements to be evaluated, you may need to give more detailed guidelines to the testers. In my case I was trying very hard NOT to break the system, and so did things like not saving flags I wasn't sure of, instead of saving them and then trying to un-apply them.

**FW**: PdBI: Specific comments which didn't quite fit into the categories above: Although I didn't do any exact benchmarking I'm quite sure that the processing with AIPS++ was slower and required much more memory than processing comparable data sets with CLIC/MIRIAD.

Logging scripts should be done by default. And the scripter
only logs things done with the task manager but not the details
parameters given to some task during interactive processing. I
principle, the logger should create a script, which might be
difficult to read, but nevertheless would execute exactly the
same commands and their options/parameter than given in the
interactive sessions.

Using field numbers instead of source names is confusing.
Especially in my case with data from two different days where
the same sources then had different field numbers.

Printing directly to the printer didn't work.

The total integration time given by the summary is wrong after
concatenating.

The help system didn't work properly: e.g. web() doesn't start
anything

ms.concatenate: it should be possible to build up a new
concatenated file without the help of a unix copy
command

A general comment:
It is very difficult to compare this aips++ testing experience
with other data reduction packages. For every package one
needs time to get to know it and all its functionalities.
During the test, obviously, I got increasingly faster in using
aips++ and also in finding out about the available
functionalities. I remember that my first encounters with
classical aips were full of frustration. This aips++ test was
certainly not that bad and I would already now prefer aips++
over aips in terms of using the package and also in some of its
functionality. On the other hand, aips++ has still to improve
its speed, functionality, simplicity of usage quite a bit to
replace MIRIAD as my mostly used interferometric data reduction
tool. And getting astronomers to switch from their old,
working software packages to aips++ will be critical for the
success of aips++.

VLA: Even more than processing the PdB data I felt that using
AIPS++ was very slow and memory hungry. Although I didn't do
any detailed benchmarking, I was reducing very similar D array
NH3 VLA data with MIRIAD the month before where the data
reductions steps were almost interactively on my machine while
with AIPS++ several times all the memory was used up almost crashing the machine (512+1024 swap) which was NEVER the case for MIRIAD!

eamples:

imgr.plotvis(type='model'); for ngc7538 uses up lots and lots of memory after filling up my whole memory and swapspace 512+512 it crashed ....

cal.correct is VERY slow

For the logger it would be helpful to know which message came from which task.

And a last general comment:

I want to stress that while there are performance problems and still lots of areas for improvements, for VLA data I prefer already now using AIPS++ instead of classical AIPS. In the current state, AIPS++ seems to be better adapter to handle VLA data then millimeter interferometer data (e.g. my PdB data set).

9  APPENDIX B: Bug Reports and Enhancement Requests

9.1  Bugs in functionality:

(1)  shadow (Reported by Chris Wilson)

    include 'iramcalibrater.g';
    shadow(msname='05may98-h121.ms', trial=F, minsep=15.0);
    shadow(msname='09may98-h121.ms', trial=F, minsep=15.0);

on 09may98 data set, it just said
    scan 4495 antenna 1 12.6811911
    scan 4496 antenna 1 12.6759066
    scan 4497 antenna 1 12.671173
    scan 4498 antenna 1 12.6669913
    scan 4499 antenna 1 12.6633624
which is what the cookbook said it should say if anything was shadowed.

** on 05may98 data, said messages like this:
    Problem with all baselines in scan 276
    Problem with shadow cal in scan 276
i.e. it doesn't seem to have found any shadowing but it has some other kind of problem. reported via email to aips2-requests@nrao.edu
George replied about 2 hours later saying shadow can't handle data when a baseline is missing. A later email the same day said it turns out that "shadowed data is NOT flagged in newly-filled data from the fits files". However, 05may98 contains no shadowed data and 09may98 is flagged correctly by shadow, so I'm still OK to proceed when ready.

(2) problem with msplot (Reported by Chris Wilson)

I had done some interactive flagging and exited saving the flags. I think I exited by typing on DONE but I'm not absolutely sure. It seemed to exit fine. I then wanted to start it up again to do some more interactive flagging and it wouldn't do it. I tried twice and got different errors. I then exited from aips++ (gracefully) and restarted it and restarted msplot from the Glish line and it seemed to start OK.

Kumar said "For some reason (which i don't understand yet) your flag table that was used to save the previous flagging was kept open although you had closed msplot."

(3) hung aips++ (Reported by Chris Wilson)

I managed to hang aips++ doing a clean of a 256x256x206 data cube. It ran for 30 minutes before I gave up and left and it was clearly slowing my machine down (window updating was slow). I still didn't have a Glish prompt when I came back the next morning, so I had to stop aips++ with kill. When I restarted aips++, the messages in the logger window suggest that the clean had actually finished, so I don't know what caused aips++ to hang. I was making a big tar file talking to a disk on a different computer and moving big files around, maybe that cause a problem with aips++? I was working in totally different subdirectories but using the same CPU.

(4) Autoflag warning messages are confusing (Reported by Deb Shepherd)

```plaintext
#include 'autoflag.g'
af:=autoflag('l1.ms');
af.setselect(timerng=['2002/01/27/05:45:47.0', '2002/01/27/07:00:00.0']);
af.run(plotscr=F);
af.done();
```
%PGPLOT, PGGRAY: foreground level = background level
%PGPLOT, PGIMAG: foreground level = background level
%PGPLOT, invalid y limits in PGSWIN: Y1 = Y2.
### Kumar says:
### "These are not important and are just warning that autoflag
### was plotting data with only one value."
Still, it degrades confidence that things are working correctly.

(5) A mistake in setting up an image crashed imager. It should handle
mistakes like these more gracefully (Reported by Deb Shepherd):

```plaintext
imgr:=imager(filename='l2.1mm.split.ms');
imgr.setdata(mode='channel', nchan=128, start=1,
step=1, spwid=[1,2,3,4], fieldid=1);
imgr.setimage(nx=256, ny=256,
cellx='0.4arcsec', celly='0.4arcsec',
stokes='I', mode='channel', nchan=1, start=10,
step=100, spwid=[1,2,3,4], fieldid=1);
imgr.clean(algorithm='clark', niter=100, gain=0.1,
model='l2.1mm.model', residual='l2.1mm.resid',
image='l2.1mm.cm', mask='', interactive=F);
```

I really needed setimage mode='mfs' since I have 2 USB and 2 LSB
spwids that I want to mfs the spwids together into a single image.
My mistake in converting a line script to continuum imaging.

Got logging errors & Glish errors. Had to exit and restart
aips++. Bug report submitted (AIPS++ defect AOCso04722)

(6) deconvolution fails (psf cannot be found) if first channel is
flagged. (Reported by Deb Shepherd)

(7) Setting different spline times at 1 and 3 mm when transferring
phase solutions from 3mm to 1mm causes the calibration to not be
determined correctly. (Found by AIPS++ group and Deb Shepherd
after test started).

(8) ms.split function did not work for Vine version of Linux (used in
Japan). Total failure at first. Kumar fixed the problem, made a
new RPM and sent to Munetake to be tested. New version still had
problems: Apparently, a parameter defined in msplot (spwid and
polarization) affected the ability to select data in split.
Bug report submitted by Deb Shepherd (AIPS++ defect AOCso04739)

(9) ms.uvlsf not working in VINE Linux AIPS++
   Bug report submitted by Deb Shepherd (AIPS++ defect AOCso04740)

(10) Raw phase transfer not working (Reported by Friedrich Wyrowski)
    For the 1mm data, tried to apply the scaled raw 3mm phases instead of the splines but couldn't get it to work with the information on the web (Although the cookbook writes this is the recommended method it only describes using the 3mm splines for this correction).

(11) Printing directly to the printer didn't work (Reported by Friedrich Wyrowski)

(12) The total integration time given by the summary is wrong after concatenating. (Reported by Friedrich Wyrowski)

(13) The help system didn't work properly: e.g. web() doesn't start anything (Reported by Friedrich Wyrowski)

(14) VLA velocity information wrong in the header, possible problem with the filler? (AIPS++ defect AOCso04559)

(15) PdBI antenna positions wrong (AIPS++ defect AOCso04329)

9.2 Problems with documentation:

(1) confusion on absolute flux calibration (Reported by Chris Wilson)

   I got confused by the plot that was produced in one of the last calibration steps, where you see amplitude versus time for all the calibrators

   cal.plotcal(tablename='ucam.3mm.temp', plottype='AMP');

   The fact that there were step functions because the flux of the various calibrators hasn't been determined yet is confusing to someone not experienced with the aips++/IRAM implementation. I forged ahead in any case and was told that the plot was OK. Maybe something to add to the cookbook i.e. a sample plot?
(2) typo in ggtau.g script in spectral line imaging (Reported by Chris Wilson)
   cal.setdata(msselect='SPECTRAL_WINDOW_ID=4')
   it should be
   cal.setdata(msselect='SPECTRAL_WINDOW_ID==4')

(3) Improve documentation about 'raw' phase transfer in the User Reference Manual. User couldn't get it to work with the information on the web (Although the cookbook writes this is the recommended method it only describes using the 3mm splines for this correction).

9.3 Enhancements identified (based on Audit comments):

(1) Job control (interrupt & abort) should stop the task, not the Glish shell.

(2) Logger window should record view-related commands

(3) Fit results should go to the main log window (so they can be cut and pasted).

(4) Shell('...') syntax needed to run host operating system commands is awkward. Should be changed.

(5) User reference manual inputs need "astronomer's definitions."

(6) Help materials should be available in PDF.

(7) Need more 'slop' for calibrators/targets to have the same spwid in VLA fast-switching data so DATA-DESC-ID does not have to be done manually in a Glish script.

(8) Provide information (documentation and history files) about how aips++ handles data weights.

(9) Provide plots of Tsys, Tau, and other M&C data if available in the dataset (need to plot this to evaluate diagnostic data and errors to make processing decisions).

(10) Bootstrapping of absolute flux density should be more flexible to incorporate decorrelation properly. Is it possible to unset fluxes set in the data files? This would be helpful for trying out different flux cal variants. I ended up having backuped the "raw" measurement sets and then copying them back to start a
fresh flux cal.

(11) Processing history associated with datasets and images should be made available.

(12) Channel selection like it is done in uvlsf should be available in other tasks.

(13) Averaging of data over time, and spectral channels should be more flexible. Averaging of data over bands should be made available.

(14) More flexible interpolation in calibration would be useful (only one default interpolation currently available).

(15) It should be possible to flag based on 'Channel 0' data and have the flags applied to the entire cube or a subset of the cube.

(16) Time variable band pass solutions should be possible.

(17) Flexible channel selection for imaging should be possible.

(18) Interactive masking should be available with MEM deconvolution.

(19) Imager should be able to input multiple datasets rather than having to concatenate them first (like MIRIAD).

(20) In uv-continuum subtraction tasks, support for offsets should be provided (if strong, point continuum source is significantly offset from the phase center).

(21) Calibration plot outputs (cal.plotcal and PdBI plots) should be controllable. Also, pgplots should be available to show PdBI calibration results rather than ps files.

(22) Provide better defaults for cube displays. I would have hoped for something more automatic. E.g. in GILDAS or MIRIAD, if I look at a cube I will see per default already more than one channel.

(23) Image display information should not be attached to an image (save/reload is not good and it is not logged).
(24) Transparent layer overlays of raster images would be nice.

(25) Seeing time-averaged display of uv data is important for low S/N data. E.g. looking at the average of all uv data in a given correlator window tells you usually immediately whether there is a line or not (if situated roughly toward the phase center). This was not possible in msplot. Could it be added?

(26) It would help immensely if visualization tasks are not bound to a measurement set. Sometimes, one wants to switch between different days or calibrators, which might have different ms files and then one has to open another msplot. This gets very complicated.

(27) For spline solutions, Phase and amplitude RMS for a given antenna solution should be given in the log window in degrees (it is currently in the fit file saved on disk but the information is important so it should go to the log window)

(30) Controlling the splines is not as easy as using polynomials. For very noisy data, I would have liked to use low order polynomials so that the fit doesn't simply follow the noise. Probably there should be both methods to do fits. I really like the MIRIAD GFIDDLE method where you can change orders of polynomials interactively, while flagging bad data at the same time.

(31) I would have liked to change the reference frequency of the data cubes to the CO 1-0 and 2-1 rest frequencies so that the velocity axis is Vlsr of the CO lines but I couldn't figure out how to do it.

(32) The On-line documentation search function on the web was sometimes difficult to navigate through lots of documentation. And in some cases the documentation referred to older aips++ versions. This should be made easier.

(33) Processing speed should be improved through benchmarks to make comparable to CLIC and MIRIAD.

(34) Source names should be used instead of field_id numbers.
## ALMA Offline Requirements - Priority 1's only

### Summary Report for all testers

**Off-line User Test:** Grades based on testing experience of functionality "only", not usability.

**Grades:** A = adequate, A/E = enhancement desired, I = inadequate, U = unable to evaluate.

**Testers:** 1 = Bergin, 2 = Wynn-Williams (VLA data), 3 = Moseley, 4 = Shepherd, 5 = Wiseman, 6 = Wynn-Williams (PDS data)

**110 requirements tested out of 273.**

<table>
<thead>
<tr>
<th>ALMA Requirement</th>
<th>VLA Data</th>
<th>PDS data</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Req #1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #28</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #48</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #51</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #52</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #53</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #54</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #55</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #59</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #60</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #61</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #65</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #66</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #69</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #71</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #78</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #79</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #84</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #85</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #87</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #88</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #90</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #93</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #97</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #99</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Req #100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Approved:** Draft, Pending, Approved, Released, Superceded, Obsolete.
<table>
<thead>
<tr>
<th>Doc #: ALMA-70.65.00.00-001-A-REP</th>
<th>Date: 2004-02-27</th>
<th>Status: Draft</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALMA Project</td>
<td>Final Report and Testing Summary</td>
<td></td>
</tr>
<tr>
<td>#</td>
<td>Criteria</td>
<td>Score</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------------------------------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>3.1</td>
<td>Sessions should be able to change all aspects of the das, including header</td>
<td>A</td>
</tr>
<tr>
<td>3.2</td>
<td>User interface should be consistent across all users in the system</td>
<td>A</td>
</tr>
<tr>
<td>3.3</td>
<td>Interface with the Archive</td>
<td>U</td>
</tr>
<tr>
<td>3.4</td>
<td>Network performance should be measured</td>
<td>A</td>
</tr>
<tr>
<td>3.5</td>
<td>Data integrity should be validated</td>
<td>A</td>
</tr>
</tbody>
</table>

**Notes:**
- A: All
- A/5: Almost all
- U: None

**Status:**
- Draft
- Pending
- Approved
- Released
- Superceded
- Obsolete

**Date:** 2004-02-27
<table>
<thead>
<tr>
<th>No.</th>
<th>Text</th>
<th>Status</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>No. 1: Refer to image of baseline data collection and model implementation</td>
<td></td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>1.2</td>
<td>No. 2: Refer to image of baseline data collection and model implementation</td>
<td></td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>1.3</td>
<td>No. 3: Refer to image of baseline data collection and model implementation</td>
<td></td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>1.4</td>
<td>No. 4: Refer to image of baseline data collection and model implementation</td>
<td></td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>1.5</td>
<td>No. 5: Refer to image of baseline data collection and model implementation</td>
<td></td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>U</td>
<td>U</td>
</tr>
</tbody>
</table>

**Notes:**
- A: Available
- B: Draft
- C: Released
- D: Superseded
- E: Obsolete
<table>
<thead>
<tr>
<th>Doc #</th>
<th>ALMA-70.65.00.00-001-A-REP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>2004-02-27</td>
</tr>
<tr>
<td>Status</td>
<td>Draft</td>
</tr>
<tr>
<td>Page</td>
<td>41 of 43</td>
</tr>
</tbody>
</table>

Table: Table of contents or a detailed list of sections.

Page 6
<table>
<thead>
<tr>
<th>#</th>
<th>Specification/Question</th>
<th>Requirement Priority</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Special Features: User Interface</td>
<td>High priority</td>
<td>Met</td>
</tr>
<tr>
<td>3.2</td>
<td>Special Features: VISI - no priority</td>
<td>Requirements</td>
<td>Met</td>
</tr>
<tr>
<td>3.3</td>
<td>Requirement must be provided by customer</td>
<td>A</td>
<td>Met</td>
</tr>
<tr>
<td>3.4</td>
<td>Requirement must be provided by customer</td>
<td>A</td>
<td>Met</td>
</tr>
<tr>
<td>3.5</td>
<td>Requirement must be provided in the user</td>
<td>A</td>
<td>Met</td>
</tr>
<tr>
<td>3.6</td>
<td>Requirement must be provided in the user</td>
<td>A</td>
<td>Met</td>
</tr>
<tr>
<td>3.7</td>
<td>Requirement must be provided in the user</td>
<td>A</td>
<td>Met</td>
</tr>
<tr>
<td>3.8</td>
<td>Requirement must be provided in the user</td>
<td>A</td>
<td>Met</td>
</tr>
<tr>
<td>3.9</td>
<td>Requirement must be provided in the user</td>
<td>A</td>
<td>Met</td>
</tr>
<tr>
<td>3.10</td>
<td>Requirement must be provided in the user</td>
<td>A</td>
<td>Met</td>
</tr>
<tr>
<td>3.11</td>
<td>Requirement must be provided in the user</td>
<td>A</td>
<td>Met</td>
</tr>
<tr>
<td>3.12</td>
<td>Requirement must be provided in the user</td>
<td>A</td>
<td>Met</td>
</tr>
</tbody>
</table>

Page 8