The Generic Instrument: IV Specifications and Development Plan

T.J. Cornwell, NRAO

Version of 1995 May 9

Contents

1 Introduction

This note follows on from previous memos by Cornwell (1995a, 1995b) on the Hamaker, Bregman, and Sault (1995) calibration formalism. Here I specify the capabilities demanded of the calibration and imaging software and lay out a development plan to cover activities for the next 12-18 months. I have worked from memo 115 in the AIPS++ User Specifications Memo series, authored by Bob Hjellming. My method has been as follows:

1. Copy those relevant sections from that memo,
2. Remove items that are irrelevant or out-dated,
3. Consolidate items into re-statements relevant to the generic instrument

I have tried hard to adhere to the original specifications, but have simplified where possible. As an example of simplification, I have tried to remove most statements that VLBI processing is special in some way. Similarly, I have assumed that mosaic processing will be standard.

Next I enumerate the priorities from NFRA (for WSRT), ATNF (for ATCA), and NRAO (for the VLA and VLBA). All of these fit within the larger set of specifications so the most important thing is the priority given to these items.

Finally, I give a development plan designed to cover work in the Synthesis area within AIPS++ over the next 12-18 months.

2 Nature of Instrumental Data

With only a few modifications, the Hjellming memo has the following requirements on the nature of the instrumental data to be addressed by AIPS++.

1. All data from the GI should be assumed to potentially (but not necessarily) involve full measurement of the electromagnetic field involving all four Stokes parameters — In hand
2. Multiple frequency bands may be simultaneously observed (e.g., for observing multiple lines simultaneously or multi-frequency synthesis), with variable numbers of channels in each band — In hand
3. Single dish and synthesis array data will be handled on an equal footing. — In hand
4. Rapid time switching of polarizations, frequencies, and pointing centers must be allowed. — OK in principle
5. Polarization measurements may be time switched if all polarization measurements are not obtained simultaneously — OK in principle
6. Data combinations for different observations may have different numbers of spectral channels and channel widths which may need to be accommodated within single data sets — In hand
7. Observations should be regarded as mosaiced by default. Mosaicing observations may have many (∼1000) pointing centers — In hand

8. Single dish observations are special cases of mosaiced data. OK in principle
9. Beam switched or multibeam systems on single dishes must be accomodated. In hand
10. If available from the on-line data archive, instrumental performance and meteorological data must be associated with instrumental data sets.
    - *Little thought has been given to this requirement as yet.*

    OK in principle
11. Meaningful error measures or estimates should be regarded as standard in an observation.
    - *This will need substantial feedback to those originating data*

    OK in principle
12. Antenna size, system temperatures, and frequency band-passes may differ widely. OK in principle
13. The integration time may vary from antenna pair to antenna pair. OK in principle
14. VLB antennas in space will require support for orbital position dependence including acceleration terms. An external emphemeris will be required. OK in principle

3 Data Correction and Calibration

3.1 General

1. Data should be selectable in terms of identification with a particular type of calibration observation
    - *This is unclear: we presume that it means that for example phase calibration observations must be distinguished from bandpass calibration observations. If so this is a far-reaching problem, extending back into the observing system.*
2. Calibration should be made as generic as possible, with telescope-specific methods kept to a minimum In hand
3. Both standard and user-defined models of data behavior should be usable in determining calibration information from data sets OK in principle
4. Instrumental behavior that affects calibration should be integrable in the calibration process through a mixture of parameterized functions and models in tabular form OK in principle
5. Data correction based upon standard and user-defined functions, with user supplied parameters, should be possible OK in principle

6. Calibration and correction of data should be reversible, with the capability to apply calibration/correction information either "on-the-fly" during processing, or "once and for all", creating new, calibrated data sets

   - *The current design has an additional column in the MeasurementSet for corrected coherences.* OK in principle

7. Calibration/correction of data should be possible from derived tables of instrumental parameters (e.g., system temperature vs. time, gain vs. elevations), with derivation of such tables from calibration observations or from on-line measurements OK in principle

8. The calibration process should include flexible averaging of calibration data and application with interpolations or weighted averaging, all under control of the user OK in principle

9. Cross-calibration from different instruments should be possible (e.g. flux scale, pointing) particular when data from different arrays are to be combined OK in principle

10. Model fitting should be possible in both the image and u-v planes, and it should be possible to use the resultant models for further calibration and self-calibration

    - *Model fitting in the u-v plane is not part of the* `MEGI` *formalism. It must be performed outside the* `MEGI` *framework.*

11. There must be simulation programs for single dish, interferometer, and mosaicing data bases for both planning and comparison of data with models - with optional error generation for thermal noise, pointing errors, primary beam errors, atmosphere, antennas surface errors, beam-switching for total power, etc. OK in principle

3.2 Interferometer Data

1. Transfer of calibration matrices from one observation to another should be possible and easy. OK in principle

2. Redundancy in data (possibly including crossing points) should be used whenever possible as an additional constraint on calibration and self-calibration Not yet clear

3. Determination of, and application of corrections for, closure errors should be possible with flexible averaging of input closure information OK in principle

4. Fringe fitting for a range of spectral channels and fringe rates should be possible by baseline, as well as globally by antenna OK in principle

5. Spectra calculation from complex summing of visibilities in each spectral channel for user-specified positions in the field of view — OK in principle

6. Interferometric pointing, baseline, and beam pattern fitting and related analysis

   - *These lie somewhat out of the* `MEGI` *framework and should be regarded as operations on Jones matrices. For example, for antenna position fitting, one would presumably use the* `MEGI` *framework to derive a G-matrix object, and then one would fit the phases in that object to find antenna positions.*

7. Application and de-application of astrometric/geodetic correction factors with complete and reversible histories — OK in principle

8. Calibration of data for effects of the ionosphere, utilizing data at multiple frequencies and/or external data on variations of electron content — OK in principle

9. Self-Calibration for non-isoplanicity must be possible

   - *The* `MEGI` *framework allows it but as far as we know, there are no demonstrated algorithms* — OK in principle

10. Determination and correction for pointing errors, and errors in beam shape, using mosaic self-calibration techniques, will be important — OK in principle

11. For spectral line sources one can do amplitude calibration with auto-correlation spectra plus calibration at one antenna — OK in principle

12. Accurate Doppler correction for each spectral channel is essential — OK in principle

13. For polarization calibration, one must be able to determine both source polarization structure and instrumental polarization ("D-term" self-calibration) — In hand

14. Full phase calibration is an iterative process involving limits set by: astrometry, geodesy, and weak source imaging/detection, therefore one needs:

    (a) very accurate geometric models, typically to at least 1/10 of a wavelength accuracy

      - *Possible eventually with the Measure system, currently being designed and implemented by Wim Brouw. Models accurate at this level will become available in late 1996.* — OK in principle

    (b) knowledge of location of the Earth's pole and UT1, both of which are generally known only after astrometric/geodetic analysis

      - *Also possible with the Measure system. Available mid-1996* — OK in principle

    (c) values of ionospheric delay as determined from measurements at simultaneous frequencies, or external measurements of ionospheric electron content — OK in principle

(d) measurement of properties of troposphere dry terms (from surface meteorological measurements) and wet terms (Kalman filtering, GPS multi-frequency satellite measurements, WVR) OK in principle

(e) instrumental delays as determined from phase calibration signals OK in principle

(f) knowledge of non-rigidity of the earth due to earth tides and atmospheric loading OK in principle

3.3 Data Editing

The requirements for data editing lie more in the domain of visualization.

1. Data display and editing should be seen as generic tools applicable to single dish, interferometer, and other forms of data OK in principle

2. Data visualization for evaluation and editing purposes should be seen as an integral, or closely coupled, aspect of the data system OK in principle

3. It should be possible to do interactive editing based upon display, with "zoom" or magnification, and menu selection of editing options OK in principle

4. Various "viewing strategies" should be available

   (a) For interferometer data, baseline by baseline display (with magnification of local areas) and interactive editing (including multiple, simultaneous baselines) using both Intensity-time-baseline displays and Intensity displays in u-v plane

   (b) Displays of spectra and spectral cubes aggregated in various ways (spectra vs. time, averaging in time, averaging of channels)

   (c) Selection of data by specifying windows in space and/or time

   (d) Selection of arbitrary cuts through data (e.g. circular, radial, or a user-defined locus) through selected data coordinates

   (e) Display of expanded data aggregates (e.g., pointing and clicking on an average multi-channel region of data to show the component spectrum

   (f) Comparison displays of generic model data (from fitted components) with observed and/or processed data, including display of data with model subtracted or divided OK in principle

5. Data editing should be reversible, with the capability to store, apply, and un-do editing information OK in principle

6. Data editing should be possible on the basis of monitor/observing log data OK in principle

7. Editing should be possible from “consistency check” information, particularly where there is redundancy or (for interferometric data) where there are crossing points in the u-v plane — OK in principle

8. It is desirable to have parameter-driven, automated flagging for large data sets — OK in principle

9. Editing must be possible based upon difference between data and models generated during self-calibration — OK in principle

10. Data editing based upon recognition of interference patterns in intensity-time-frequency data is very important, particularly for low frequency observations — OK in principle

4 Imaging and Image Processing

In this section we consider the formation of images from edited, calibrated data. While this is mainly image computation and deconvolution, it must be remembered, that for the user, imaging and image deconvolution is an integral part of the process of data inspection/editing, calibration, imaging, self-calibration, data/image display, spectrum/time/image analysis, and production of hard copy for publication purpose. This process must be well integrated for the convenience of the user. It should be possible to easily “mix-and-match” self-calibration, data transformation, and de-convolution “tools”, for example, using CLEAN to deconvolve in the early stages, and maximum entropy later on when CLEAN begins to be less useful. This is related to the need to make self-calibration use a generic model, which could be a table of CLEAN-components, a table of Gaussian components, or an image.

4.1 Image and Spectral Image Formation

1. Image construction using u-v data sets must be possible with a range of capabilities

   (a) Computation of “dirty” images and point spread functions by 2-D FFT of selected data with user control of data selection, gridding algorithm and its parameters, and image parameters (image size, cell sizes, polarization) — In hand

   (b) Flexible computation of data cubes where the third axis is frequency/velocity or time — OK in principle

   (c) Simultaneous, multiple field imaging — OK in principle

   (d) Direct Fourier transform imaging of arbitrary (and usually small) size fields — In hand

(e) Imaging after subtraction for sources — OK in principle

(f) Imaging of spectral line data sets with continuum subtraction based upon continuum data, or continuum models — OK in principle

(g) Estimation and input of zero-spacing flux density and appropriate weighting — OK in principle

2. Mosaic image construction using mixture of u-v data sets and single dish data for multiple antenna pointing centers

   (a) Linear combination of pre-deconvolved images, weighting determined by primary beam — OK in principle

   (b) Linear mosaic algorithm with linear deconvolution (MOSLIN in SDE) — OK in principle

   (c) Non-linear (MEM-based) mosaic algorithm (VTESS, UTESS in AIPS, mosaic in SDE) — OK in principle

   (d) Cross-calibration (enforced consistency) between data taken with different instruments (flux scale, pointing) — OK in principle

   (e) Pointing self-calibration to determine corrections for both single dish and visibility data — OK in principle

   (f) Non-coplanar baselines mosaicing allowing for sky curvature — OK in principle

   (g) Self-calibration and editing of all pointings in one processing step — OK in principle

   (h) Capability to determine the primary beam(s) from a mosaic image and its related data sets — OK in principle

   (i) Ability to deal with any primary beams in different forms (analytic 1- and 2-D, tabular), including user modification of primary beam models — OK in principle

3. Imaging using multiple-frequency data sets and a user-defined model for spectral combination "rules" must be possible — OK in principle

4. Imaging computation should generally take multiple data sets where this makes sense — OK in principle

5. Imaging data selection should flexibly allow use of data sub-sets, with data selection based upon time, antenna, frequency, and ranges of other data (including monitor data) — OK in principle

6. Non-coplanar baselines imaging (dragon in SDE) — OK in principle

7. Imaging wide fields large than the isoplanatic region — OK in principle

8. Near field imaging of nearby objects like comets and asteroids — OK in principle

9. Fringe-rate imaging — OK in principle

5 Priorities for development

While the previous sections reflect the specifications for AIPS++ functionality in the area of synthesis processing, the priorities for development are not given. In this section, I discuss these priorities.

5.1 Completion of MEGI design

Cornwell and Wieringa (1996) describe the design of the MEGI. The following have yet to be completed:

1. Scalar version for *e.g.* RR alone or RR and LL only.
2. Full persistence of objects
3. Optimization for speed, including gridding and Fourier transformation
4. Cross-calibration

5.2 NFRA

NFRA wants to use AIPS++ for data analysis for the new WSRT on-line system TMS. TMS is expected to debut in August 1996. NFRA needs a commitment from AIPS++ to support such use of AIPS++. As outlined by Jan Noordam, the priorities are:

1. AIPS++ port to HP/UX
2. Fill to AIPS++ Synthesis MeasurementSet from WSRT data format
3. Simple calibration, editing and imaging of WSRT data, controllable from TMS
4. Polarized sky models including XX,YY,XY and YX as well as the standard I,Q,U,V
5. Parametrized source components. Initially only for existing calibrator models.
6. Subtraction of known sources from coherences
7. Simple data statistics: averages, rms, etc.
8. Visualization of coherence data

5.3 ATCA

Mark Wieringa outlined priorities for ATCA calibration :

1. A Solver for bandpass and gain using the parallel hand correlations.
2. A Solver for polarization leakage, gain and optionally source polarization using all four correlations.
3. Selfcal Solvers using either single Stokes (I) correlation data or multiple correlations.
4. A Corrector for each of bandpass, leakage and antenna/i.f.-gain and a versioning scheme for either the Correctors or their underlying tables.

There is no specific “drop-dead” date attached to these priorities since Miriad currently can be used.

5.4 BIMA

Peter Teuben (1996) has outlined the special needs of BIMA.

1. Support for time-sliced fashion polarization measurements, measuring circular polarization (LR and RL) with a quarter-wave plate. See also Wright (1995a) for a discussion on some of the possibilities.
2. Deconvolution of mosaiced fields, including pointing corrections. Also adding single dish data to interferometry data.
3. VLBI: BIMA regularly participates in mm VLBI experiments. The phased array data are currently processed offline using standard VLBI techniques (Mark-xxx, ref. xxx), and whenever AIPS++ will provide VLBI data processing, BIMA should be able to use them without any major problems.
4. Heterogenous array elements.
5. Unusual correllator modes: The correlator can be configured in many modes, and produces DSB data with a small (¡ 8) number of windows with different settings of the IF. An interesting method to calibrate DSB data is to use the generally much slower varying gain ratio (phase difference and amplitude ratio).
6. Offline phase corrections, using total power measurements (see Wright 1995b), examplify one of the many ways in which calibration needs to be flexible.

MIRIAD employs a very general visibility file format, where correlations (both cross- and auto) are tagged with a rich set of (name based) variables. Variables can be multi-dimensional of any of the basic types (variables can change dimension in a dataset, in principle even type). Obviously, like in the FITS community, these variables need to be registered and their meaning clarified. Currently MIRIAD knows about 95 variables. Although it would be ideal that the telescope data be directly written in native AIPS++ format, for the foreseeable future, a conversion program will be used.

5.5 VLA

A mail message from Michael Rupen dated Sept 19, 1995, gives the following priorities for VLA software development.

1. Interpolating bandpass solutions.
2. Ionospheric corrections from GPS.
3. D-term self-calibration (with and without time-variability).
4. Mosaicing.
5. High dynamic range imaging (Briggs' NNLS algorithm)
6. Automated flagging.

In general, NRAO will use whatever system is appropriate to acheive a given functionality. In some cases, AIPS is still the chosen route, whereas in others AIPS++ is preferred. The first and second items are being addressed within AIPS.

5.6 VLBI

Tony Beasley has summarized the special requirements for VLBI:

1. High precision quantities
2. Version typing to allow tracking of e.g. interferometer models used
3. Multi-file datasets (greater than 2 Gbyte and perhaps spread over many disks)
4. Variable and possibly unequal integration times
5. Provision for tied arrays
6. Diversity of antenna mounts
7. "In the beam" phase referencing

In addition, tasks for the following will be needed:

1. Data readers for various formats
2. Sophisticated model for correlator effects e.g. decorrelation, state-count corrections.
3. Absolute amplitude calibration is more important than in connected element interferometry. There are various approachs, all of which must be supported.
4. Fringe fitting, both antenna and baseline-based.
5. External data calibration
6. Phase-cal information
7. Polarization calibration, including all known calibration algorithms
8. Velocity correction for fringe-rotation
9. Pulsar binning and gating
10. Source fitting

6 Development plan

Here we cover the parts of the development plan appropriate to synthesis imaging. The goals of this development plan are to address pressing needs of the consortium partners and to push forward into new application areas. It is expected that this development plan will be revised from time to time to incorporate, for example, VLBI processing when appropriate.

AIPS++ port to g++	June 96	NFRA
Final version of WSRT Filler	July 96	Olnon
Spectral support	July 96	Wieringa
Cross calibration and editing	July 96	Cornwell
ComponentSkyModel	July 96	Marson
TMS commissioning	Aug 96	de Vos
VLA 90cm imaging	Oct 96	Cornwell
Mosaicing	Dec 96	Cornwell

Some notes:

Mosaicing support is a way to test the framework for inclusion of image plane effects.

Spectral support means that the MEGI framework must be extended to allow multiple spectral channels. This is straightforward,

VLA 90cm imaging will be accomplished using a program similar in principle to dragon in SDE.

References

Briggs, D.S., 1995, Ph.D. thesis, New Mexico Institute of Mining and Technology

Cornwell, T.J., 1995, AIPS++ Implementation Note 183.

Cornwell, T.J., 1995, AIPS++ Implementation Note 184.

Cornwell, T.J., and Wieringa, M.H., 1996, AIPS++ Implementation Note 189.

Hamaker, J.P., Bregman, J.D., and Sault, R.J., 1995, *Understanding radio polarimetry: I Mathematical foundations*, submitted to A&A.

Teuben, P, 1996, BIMA calibration requirements for AIPS++