VLASS Project Memo #12

VLASS Wideband Stokes I Single Epoch Mosaic Imaging Test and Implementation Plan

Claire J. Chandler and Steven T. Myers (NRAO)

November 14, 2018

1. Introduction

The Algorithm R&D Group recently produced a very thorough memo describing the remaining issues with the 'awproject' and 'mosaic' gridders in CASA task 'tclean', along with an analysis of problems with the VLASS data, and computing estimates for wideband mosaic imaging: see https://safe.nrao.edu/wiki/pub/Software/Algorithms/WebHome/ARDG_VLASS_Imaging_Report_v2.pdf. As a result of that memo, decisions now have to be made regarding the potential use of these imaging algorithms for VLASS, based on considerations of scientific accuracy and compute costs. This memo describes a path toward for making these decisions, comprising a summary of outstanding bugs that needs to be addressed in the various algorithms, work that needs to be done to improve the intrinsic quality of the VLASS data, tests that need to be carried out in order to finalize the compute requirements, and a description of the required modifications to the Single Epoch imaging pipeline.

2. Algorithms and Software

This section presents the blocking items and their current status based on Chapter 3 of the ARDG memo.

2.1 CASA gridders

The ARDG memo demonstrated that the accuracy of the spectral index derived from the 'mtmfs' deconvolution is very sensitive to the pointing accuracy of individual antennas and the correct handling of the primary beam. Based on the information in the memo, we identify the following areas of development as being blockers for VLASS wideband mosaics, for each of the gridders:

'awproject' gridder:

1. Ability to read VLA pointing tables (this currently has a bug, and it would be good to unify the method with that used by the mosaic gridder)
2. Ability to apply antenna-dependent pointing offset corrections
3. Resolve PB projection issues (CAS-7727)

‘mosaic’ gridder:

1. Fix the use of the conjbeams parameter (CAS-11902)
2. Ability to read VLA pointing tables and apply antenna-dependent pointing offset corrections (CAS-11191)
3. Use of PB model with sidelobes (CAS-9968 with vptable)
4. Resolve PB projection issues (CAS-7727)

The need for w-term corrections was demonstrated for the BnA dataset used for the ARDG tests. In order to minimize the errors in source positions and smearing introduced by large w-terms the VLASS Project will apply a uvrange selection to the input data for the BnA hybrid. Other developments for both gridders have been suggested, including the introduction of w-term corrections into the mosaic gridder, and the ability to limit the largest w to be used in the calculation of the convolution function in the awproject gridder. These latter developments should be pursued, and the VLASS Team will help with testing as needed. But based on the timing data below, further performance improvements may be needed before the use of w-projection can become a reality all-sky for VLASS (note, however, that it may be useful, and even critical, for a small fraction of the sky with bright sources).

2.2 Addressing phase/pointing offsets in the data

The ARDG memo identified problems with both phase and amplitudes of the input data. The phase problems most likely arise from the well-known ‘delay clunking’ associated with the VLA correlator, and bright sources where this may limit dynamic range can be self-calibrated. The amplitude errors have been demonstrated to arise from a bug in the online system for antennas with old-style ACUs. This bug has been fixed for VLASS1.2 and following, but corrections are needed for VLASS1.1, and lead to some of the CASA development requirements described in section 2.1, in addition to work from the VLASS Project to correct the pointing tables for the antennas with old ACUs for use by tclean. The correction of VLASS1.1 data is described in Section 3.

2.3 Pipeline modifications

The ARDG memo suggested modifying the way in which the CASA task ‘statwt’ is used in the VLASS pipeline to allow it to use the residual (data-model) for deriving the data weights. Minor changes in the pipeline task hifv_statwt may be needed to support this suggestion, and an initial test shows that considerable improvement in the mosaic PB (see Figure 1). Note that this same (data-model) residual can be used for improved RFI subtraction in the imaging pipeline as well.

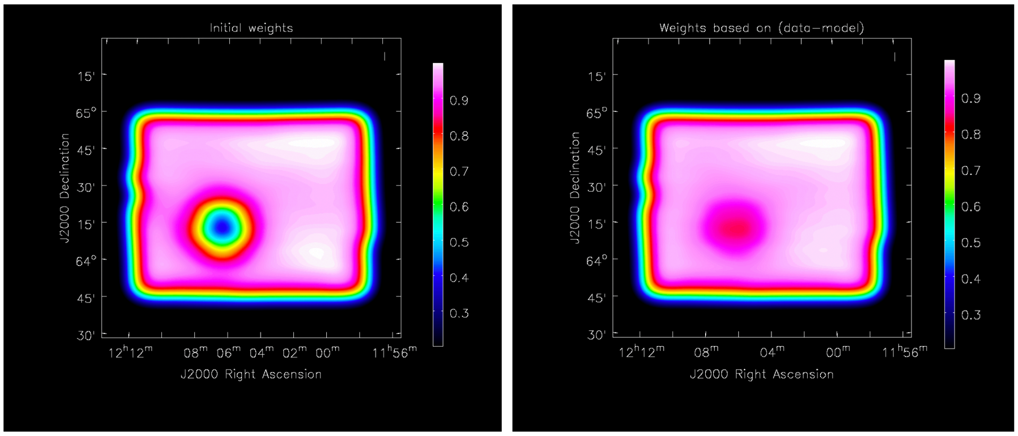

Figure 1: The mosaic PB with the original weights derived by the VLASS calibration pipeline for a bright source in the GOODSN field (left) and for weights derived using the residual after subtracting a model based on a Quick Look image (right). The minimum in the PB around the bright source increases from 0.40 to 0.84 with the new weights.

3. Correcting problems with VLASS1.1 data

The primary problem with the VLASS1.1 data originated from a bug in the online system that caused antennas with old ACUs to be mis-pointed during an OTF row, as described below based on an email from Bryan Butler:

> …During VLASS1.1 observing, there was an error in the on-line software system that caused all antennas with old ACUs to be pointed ahead of where they should have been along the OTF line, in both azimuth and elevation. The description of the dates and details of this pointing error is as follows. In August 2017, in an attempt to reconcile where antennas were being commanded to be pointed with where they were actually pointed, along with where the new SDM Pointing table entries indicated that they were pointed, a “look-ahead” was put into the MIB software for the old ACUs (not needed for the new ACUs because of the different nature by which the MIB controls them vs. the old ones). Because new pointing polynomials can only be sent to the old ACUs every 100 msec, and actual positions are read back at twice that rate, the proper magnitude of this look-ahead was 50 msec (to project to the middle of the period for which the pointing polynomial was valid, and the time of the next read back). Unfortunately an error was made in the MIB software, and 500 msec was used for the look-ahead time instead of 50 msec. This is the magnitude of the time offset in azimuth. A second bug was introduced where a factor of 2 was multiplied into that look-ahead time for elevation, so that time offset is 1 sec. The error therefore, in time, for the two axes is 450 and 950 msec, for azimuth and elevation respectively. To get the magnitude of the error in angle on the sky, that time offset must be multiplied by the position rate for each axis. These two rates will not be constant for all of VLASS1.1, because the OTF lines are at fixed declination, and so the azimuth and elevation rates will depend in detail on position in the sky, and time of observation. That rate should be able to be calculated from the SDM Pointing table itself, from the commanded positions column (by differencing subsequent entries).

Because of the way the pointing information is written into the SDM for the VLA, the pointing tables are incorrect and need to be modified. The following antennas had old ACUs during the following time periods of VLASS1.1:

Date range	Antennas with old ACUs
9/7/17-10/23/17	ea03, ea04, ea05, ea06, ea07, ea09, ea10, ea11, ea12, ea13, ea15, ea16, ea18, ea19, ea20, ea22, ea23, ea24, ea25, ea26, ea27, ea28
10/23/17-2/7/18	ea03, ea04, ea05, ea06, ea07, ea09, ea10, ea11, ea12, ea13, ea15, ea16, ea18, ea19, ea20, ea22, ea23, ea24, ea25, ea26, ea27
2/7/18- 2/20/18	ea03, ea04, ea05, ea06, ea09, ea10, ea11, ea12, ea13, ea15, ea16, ea18, ea19, ea20, ea22, ea23, ea24, ea25, ea26, ea27

After the pointing tables have been corrected, tclean needs to be able to use them during the imaging (CAS-11191). When this functionality is in place the imaging tests carried out in the ARDG memo using the BnA dataset will be repeated (initially using the mosaic gridder, which is closer to being able to use the pointing tables than the awproject gridder). These data will also be used to establish limits on the accuracy of the spectral index achievable using OTF.

4. Choice of gridder and compute requirements

4.1 Choice of gridder for VLASS Stokes I SE imaging

The compute requirements are highly dependent on the choice of gridder. At this time we face the trade-off of accuracy (through application of w-term corrections during the imaging) and compute resources required. The ARDG memo demonstrated that the use of w-terms increases the run-time of the imaging by a factor of 50 over using no w-terms. This factor of 50 makes the imaging of VLASS data impractical using NRAO computing hardware, although resources may exist externally that could support this load. For expediency, particularly the additional software development that would be needed to run the imaging on external resources while interacting with the VLASS Manager, we will start the processing of SE images for the majority of the sky without w-terms. The impact of this choice will be a dynamic range limit for images containing bright sources. Areas of sky affected by this dynamic range limit can be processed using w-terms as needed; it is therefore critical to understand what this dynamic range limit is, so that an evaluation of the fraction of sky that might need to be imaged with w-terms can be made.

> *Tests needed:* Using conjbeams=True, compare use of wprojplanes=1,32,128, plus wprojplanes=32 with WTerm_WFUDGE=16 to evaluate source brightness where images are limited by dynamic range rather than thermal noise. These tests must be carried out with data not affected by the pointing problem described above, so that the w-term issues can be properly isolated and characterized (i.e., pilot data and simulations).

The ARDG memo also demonstrated that using conjbeams=True provides accurate spectral indices, and that this option can be used for wideband VLASS SE images at the cost of reduced sensitivity. At this time both the awproject and mosaic gridders can be invoked with conjbeams=True (modulo possible issues noted in CAS-11902 for the mosaic gridder). Thus the accuracy of the spectral index may depend on the accuracy of the different primary beams used by those gridders, including sidelobes.

Tests needed: Using conjbeams=True, compare use of the mosaic gridder, using both the standard beam and a vptable beam with sidelobes, with the awproject gridder using wprojplanes=1. Specifically, the rate of divergences and values of source spectral indices need to be evaluated. These tests must be carried out with data not affected by the pointing problem described above, so that the primary beam issues can be properly isolated and characterized (i.e., pilot data and simulations).

Finally, the effectiveness of applying the pointing correction for those antennas noted in Section 3 above needs to be evaluated.

Tests needed: Using both conjbeams=True and conjbeams=False, compare imaging with and without applying the pointing table using the mosaic gridder. Specifically, the level of residual amplitude errors, residual PSF artifacts, and accuracy of derived spectral indices needs to be evaluated. These tests should use fields observed in both the pilot and VLASS1.1 (e.g., GOODSN), and the T14t05 field that was observed twice in VLASS1.1 because of the Moon and which showed the original flux density discrepancy as a function of declination. In addition, the dataset used for the ARDG analysis, in which the phase calibrator was covered within the OTF image, can be used to verify that application of the pointing corrections results in accurate flux densities and spectral index.

4.2 Compute requirement tests

Having made the decision to use the mosaic gridder where possible, and the awproject gridder for those parts of the sky where bright sources exhibit dynamic range limitations due to the lack of w-term corrections with the mosaic gridder, we need to do the following to evaluate the computing requirements:

1. Determine the fraction of sky for which the mosaic gridder can be used;
2. Evaluate the run-time of major-minor cycles in tclean for both wideband and coarse cube imaging;
3. Calculate the associated compute requirements.

All test results are being tracked via a Google spreadsheet at
https://docs.google.com/spreadsheets/d/1spWRW6SUVyJAivGPYE_2qVzlso06p3y2GcvrMwprR_Y/edit?usp=sharing

5. Single Epoch Imaging Pipeline

The SE wideband Stokes I imaging pipeline will need to be updated according to the results of the tests described in Section 4. At this time we can assert that the following will be needed in the SE wideband imaging pipeline based on the ARDG recommendations listed in Section 2.3:

1. The ability to determine the data weights based on (data-model), as described in CASA requirements ticket CASR-361, needs to be added to the pipeline task hifv_statwt;
2. Additional RFI flagging based on (data-model), as described in CASA requirements ticket CASR-362, should be included in the SE imaging pipeline.
3. Further modifications of the calls to tclean may be needed.

6. Schedule

W-term tests:

> As of 11/13/18 the awproject tests using different numbers of w-planes with VLASS pilot data are all complete, and need to be analyzed. This will be completed by 11/16/18. Refinement of the results using simulations will take until 12/7/18.

PB tests:

> As of 11/13/18 the initial set of tests using different primary beams with the mosaic and awproject gridders, based on VLASS pilot data, are all complete, and indicate problems possibly related to the issues reported in CAS-11902. They will need to be repeated when these issues have been resolved.

Pointing correction tests:

> As of 11/13/18 the VLASS team is waiting on a CASA build with the ability to apply pointing corrections (CAS-11191). This is currently under development for CASA 5.5.0, for which the feature freeze is 12/15/18. The VLASS team will be working with the developer (K. Golap) on testing and debugging as soon as a test build is available.

Compute requirement calculations:

> While the computing requirements are ideally determined using the final versions of the CASA tclean code that will be used for the SE imaging, we can make initial estimates using the existing code, once the w-term analysis is complete. These initial estimates can be completed by 12/31/18. They will have to be modified when the performance impact of applying the pointing corrections has been evaluated.

7. Summary

This memo identifies development priorities for the mosaic and awproject gridders in the CASA task tclean (as described in JIRA tickets CAS-11902, CAS-11191, CAS-9968, and CAS-7727) in order to support VLASS SE Stokes I imaging, many of which are already completed or in active development for CASA 5.5.0. It also describes a series of tests needed to determine when w-projection is needed, to determine which primary beams are most effective at minimizing divergences, to evaluate whether the application of pointing corrections improves the images, and to determine the computing requirements for SE imaging. Modifications to the SE wideband Stokes I imaging pipeline will be needed, based on the outcome of these tests.

Note that additional tests and modifications to the imaging pipeline will be needed to support the SE coarse cubes, which will be addressed separately.