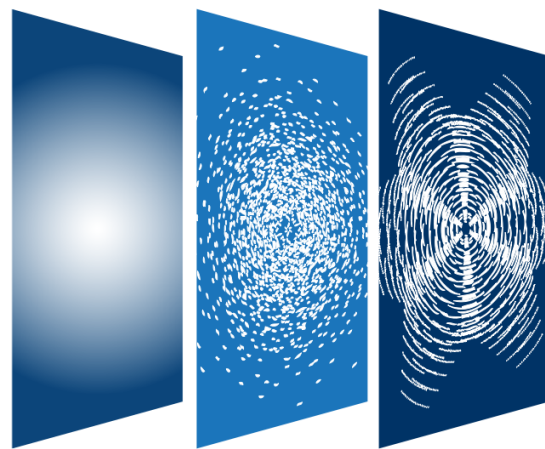


CASA User Reference & Cookbook

Release 4.5.2



CASA

Common Astronomy
Software Applications

Version: February 10, 2016

Permission is granted to make and distribute verbatim copies of this manual provided the copyright notice and this permission notice are preserved on all copies.

Permission is granted to copy and distribute modified versions of this manual under the conditions for verbatim copying, provided that the entire resulting derived work is distributed under the terms of a permission notice identical to this one.

Permission is granted to copy and distribute translations of this manual into another language, under the above conditions for modified versions, except that this permission notice may be stated in a translation approved by AUI.

- CASA homepage: <http://casa.nrao.edu>
- For help, comments, or corrections, please visit the NRAO helpdesk: <http://help.nrao.edu> or <http://help.almascience.org> for ALMA related questions.
- Data reduction guides are available at <http://casaguides.nrao.edu>
- For news, updates, critical bugs, subscribe to our CASA mailing lists: http://casa.nrao.edu/mail_list.shtml
- Join the NRAO Users community on <http://science.nrao.edu/forums>

CASA Synthesis & Single Dish Reduction Reference Manual & Cookbook

Editor: Jürgen Ott – CASA Project Scientist

Jeff Kern – CASA Project Manager

<http://casa.nrao.edu>

CASA SYNTHESIS & SINGLE DISH REDUCTION
REFERENCE MANUAL & COOKBOOK,
Version February 10, 2016,
©2015 National Radio Astronomy Observatory

The National Radio Astronomy Observatory is a facility of the National Science Foundation
operated under cooperative agreement by Associated Universities, Inc.

Contents

1	Introduction	28
1.0.1	Reference for Publications	30
1.1	About This Release	30
1.2	Obtaining CASA	31
1.2.1	What's New in Release 4.5.2	31
1.3	CASA Basics — Information for First-Time Users	33
1.3.1	Before Starting CASA	33
1.3.1.1	Environment Variables	34
1.3.1.2	Where is CASA?	35
1.3.2	Starting CASA	35
1.3.3	Ending CASA	36
1.3.4	What happens if something goes wrong?	36
1.3.5	Aborting CASA execution	36
1.3.6	What happens if CASA crashes?	37
1.3.7	Python Basics for CASA	37
1.3.7.1	Variables	37
1.3.7.2	Lists and Ranges	38
1.3.7.3	Indexes	38
1.3.7.4	Indentation	39
1.3.7.5	System shell access	39
1.3.7.6	Executing Python scripts	39
1.3.8	Getting Help in CASA	40
1.3.8.1	TAB key	40
1.3.8.2	<code>help <taskname></code>	40
1.3.8.3	<code>help</code> and <code>PAGER</code>	43
1.3.8.4	<code>help par.<parameter></code>	43
1.3.8.5	Python <code>help</code>	44
1.4	Tasks and Tools in CASA	44
1.4.1	What Tasks are Available?	45
1.4.2	Running Tasks and Tools	51
1.4.2.1	Aborting Synchronous Tasks	52
1.4.3	Getting Return Values	52
1.4.4	Running Tasks Asynchronously	53
1.4.5	Setting Parameters and Invoking Tasks	53

1.4.5.1	The scope of parameters in CASA	55
1.4.5.2	The <code>default</code> Command	56
1.4.5.3	The <code>go</code> Command	57
1.4.5.4	The <code>inp</code> Command	58
1.4.5.5	The <code>saveinputs</code> Command	61
1.4.5.6	The <code>tget</code> Command	62
1.4.5.7	The <code>tput</code> Command	63
1.4.5.8	The <code>.last</code> file	63
1.4.6	Tools in CASA	64
1.5	Getting the most out of CASA	65
1.5.1	Your command line history	65
1.5.2	Logging your session	65
1.5.2.1	Startup options for the <code>logger</code>	68
1.5.2.2	Setting priority levels in the <code>logger</code>	69
1.5.3	Where are my data in CASA?	70
1.5.3.1	How do I get rid of my data in CASA?	70
1.5.4	What's in my data?	71
1.5.5	Data Selection in CASA	72
1.6	From Loading Data to Images	72
1.6.1	Loading Data into CASA	72
1.6.1.1	VLA: Filling data from VLA archive format	74
1.6.1.2	Filling data from UVFITS format	74
1.6.1.3	Loading FITS images	74
1.6.1.4	Concatenation of multiple MS	74
1.6.2	Data Examination, Editing, and Flagging	74
1.6.2.1	Interactive X-Y Plotting and Flagging	75
1.6.2.2	Flag the Data Non-interactively	75
1.6.2.3	Viewing and Flagging the MS	75
1.6.3	Calibration	75
1.6.3.1	Prior Calibration	76
1.6.3.2	Bandpass Calibration	77
1.6.3.3	Gain Calibration	77
1.6.3.4	Polarization Calibration	77
1.6.3.5	Examining Calibration Solutions	78
1.6.3.6	Bootstrapping Flux Calibration	78
1.6.3.7	Correcting the Data	78
1.6.3.8	Splitting the Data	78
1.6.3.9	UV Continuum subtraction	79
1.6.3.10	Transforming the Data to a new frame	79
1.6.4	Synthesis Imaging	79
1.6.4.1	Cleaning a single-field image or a mosaic	79
1.6.4.2	Feathering in a Single-Dish image	80
1.6.5	Self Calibration	80
1.6.6	Data and Image Analysis	80
1.6.6.1	What's in an image?	81

1.6.6.2	Image statistics	81
1.6.6.3	Image values	81
1.6.6.4	Moments of an image cube	81
1.6.6.5	Image math	81
1.6.6.6	Regridding an Image	81
1.6.6.7	Displaying Images	82
1.6.7	Getting data and images out of CASA	82
2	Visibility Data Import, Export, and Selection	83
2.1	CASA Measurement Sets	84
2.1.1	Under the Hood: Structure of the Measurement Set	84
2.2	Data Import and Export	88
2.2.1	ALMA: Filling of Science Data Model (ASDM) data	88
2.2.1.1	Import of ASDM data with option <i>lazy=True</i>	93
2.2.2	Jansky VLA: Filling of Science Data Model (ASDM) data	94
2.2.3	VLA: Filling data from archive format (<code>importvla</code>)	96
2.2.3.1	Parameter <code>applytsys</code>	97
2.2.3.2	Parameter <code>bandname</code>	98
2.2.3.3	Parameter <code>frequencytol</code>	98
2.2.3.4	Parameter <code>project</code>	98
2.2.3.5	Parameters <code>starttime</code> and <code>stoptime</code>	99
2.2.3.6	Parameter <code>autocorr</code>	99
2.2.3.7	Parameter <code>antnamescheme</code>	99
2.2.3.8	Parameter <code>evlabands</code>	99
2.2.4	Import MIRIAD visibilities (<code>importmiriad</code>)	99
2.2.5	UVFITS Import and Export	100
2.2.5.1	Import using <code>importuvfits</code>	100
2.2.5.2	Import using <code>importfitsidi</code>	100
2.2.5.3	Export using <code>exportuvfits</code>	101
2.2.6	Handling Measurement Set metadata and data	102
2.2.7	Summarizing your MS (<code>listobs</code>)	103
2.2.8	MMS summary (<code>listpartition</code>)	105
2.2.9	Listing MS data (<code>listvis</code>)	105
2.2.10	Listing and manipulating MS metadata (<code>vishead</code>)	106
2.2.11	MS statistics (<code>visstat</code>)	108
2.2.12	Concatenating multiple datasets (<code>concat</code>)	109
2.3	Data Selection	111
2.3.1	General selection syntax	112
2.3.1.1	String Matching	112
2.3.2	The <code>field</code> Parameter	113
2.3.3	The <code>spw</code> Parameter	114
2.3.3.1	Channel selection in the <code>spw</code> parameter	115
2.3.4	The <code>selectdata</code> Parameters	116
2.3.4.1	The <code>antenna</code> Parameter	116
2.3.4.2	The <code>scan</code> Parameter	118

2.3.4.3	The <code>timerange</code> Parameter	118
2.3.4.4	The <code>uvrange</code> Parameter	119
2.3.4.5	The <code>correlation</code> Parameter	120
2.3.4.6	The <code>intent</code> Parameter	120
2.3.4.7	The <code>observation</code> Parameter	120
2.3.4.8	The <code>msselect</code> Parameter	121
3	Data Examination and Editing	122
3.1	Plotting and Flagging Visibility Data in CASA	122
3.2	Managing flag versions with <code>flagmanager</code>	123
3.3	X-Y Plotting and Editing of the Data	124
3.3.1	MS Plotting and Editing using <code>plotms</code>	124
3.3.1.1	Loading and Selecting Data	127
3.3.1.2	A Brief Note Regarding <code>plotms</code> Memory Usage	128
3.3.1.3	Plot Axes	128
3.3.1.4	Calibration Library	131
3.3.1.5	Tools	132
3.3.1.6	Interactive Flagging in <code>plotms</code>	133
3.3.1.7	Averaging Data	134
3.3.1.8	Plot Symbols	136
3.3.1.9	Summarizing Data	136
3.3.1.10	Defining Frequency and Velocity	137
3.3.1.11	Shifting the Phase Center	138
3.3.1.12	Plot Ranges	138
3.3.1.13	Plot Labels	138
3.3.1.14	Grid Lines	139
3.3.1.15	Legend	139
3.3.1.16	The Options Tab	139
3.3.1.17	Iteration	139
3.3.1.18	Overplotting	140
3.3.1.19	Plotting Multiple Data Sets	140
3.3.1.20	Saving your plot	142
3.3.1.21	Exiting <code>plotms</code>	142
3.3.2	Plotting and Editing using <code>plotxy</code>	143
3.3.2.1	GUI Plot Control	146
3.3.2.2	The <code>selectplot</code> Parameters	147
3.3.2.3	The <code>iteration</code> parameter	148
3.3.2.4	The <code>overplot</code> parameter	148
3.3.2.5	The <code>plotrange</code> parameter	148
3.3.2.6	The <code>plotsymbol</code> parameter	149
3.3.2.7	The <code>showflags</code> parameter	150
3.3.2.8	The <code>subplot</code> parameter	151
3.3.2.9	Averaging in <code>plotxy</code>	151
3.3.2.10	Interactive Flagging in <code>plotxy</code>	153
3.3.2.11	Flag extension in <code>plotxy</code>	153

3.3.2.12	Setting rest frequencies in <code>plotxy</code>	155
3.3.2.13	Printing from <code>plotxy</code>	156
3.3.2.14	Exiting <code>plotxy</code>	157
3.3.2.15	Example session using <code>plotxy</code>	157
3.3.3	Plotting antenna positions using <code>plotants</code>	160
3.3.4	Plotting uv-coverages <code>plotuv</code>	160
3.4	Data Flagging using <code>flagdata</code>	161
3.4.1	The <code>action</code> parameter	162
3.4.2	Flagging Modes	163
3.4.2.1	Manual Flag/Unflag	163
3.4.2.2	List	164
3.4.2.3	Clip	165
3.4.2.4	Shadow	165
3.4.2.5	Quack	166
3.4.2.6	Elevation	166
3.4.2.7	Tfcrop	167
3.4.2.8	Rflag	169
3.4.2.9	Extend	171
3.4.2.10	Unflag	173
3.4.2.11	Summary	174
3.5	Command-based flagging using <code>flagcmd</code>	174
3.5.1	Input modes <code>inpmode</code>	175
3.5.1.1	Input mode <code>'table'</code>	175
3.5.1.2	Input flag mode <code>'list'</code>	176
3.5.1.3	Input flag mode <code>'xml'</code>	176
3.5.2	Operation types <code>action</code>	177
3.5.2.1	Apply flags — <code>optype</code> option <code>'apply'</code>	177
3.5.2.2	Unapply flags — <code>action</code> option <code>'unapply'</code>	178
3.5.2.3	List flags — <code>action='list'</code>	178
3.5.2.4	Plot flags — <code>action='plot'</code>	179
3.5.2.5	Clear flags — <code>action='clear'</code>	179
3.5.2.6	Extract Flag Commands— <code>action='extract'</code>	179
3.5.3	Flagging command syntax	180
3.6	Browse the Data	182
4	Synthesis Calibration	186
4.1	Calibration Tasks	186
4.2	The Calibration Process — Outline and Philosophy	188
4.2.1	The Philosophy of Calibration in CASA	191
4.2.2	Keeping Track of Calibration Tables	191
4.2.3	The Calibration of traditional VLA data in CASA	192
4.2.4	Loading Jansky VLA data in CASA	193
4.3	Preparing for Calibration	193
4.3.1	Weight initialization and <code>WEIGHT_SPECTRUM</code>	194
4.3.2	System Temperature and Switched-Power Corrections	195

4.3.3	Antenna Gain-Elevation Curve Calibration	196
4.3.4	Atmospheric Optical Depth Correction	196
4.3.4.1	Determining opacity corrections for <i>modern</i> VLA data	197
4.3.4.2	Determining opacity corrections for VLA data	198
4.3.5	Setting the Flux Density Scale using (<code>setjy</code>)	198
4.3.5.1	Using Calibration Models for Resolved Sources	203
4.3.6	Correction for delay and antenna position offsets using <code>gencal</code>	205
4.3.7	Applying Jansky VLA switched power or ALMA Tsys using <code>gencal</code>	209
4.3.8	Generate a gain table based on Water Vapor Radiometer data <code>wvrgcal</code>	209
4.3.8.1	Statistical parameters shown in the logger output of <code>wvrgcal</code>	211
4.3.8.2	Antenna position calculation	212
4.3.8.3	Spectral window selection	213
4.3.9	Ionospheric corrections	213
4.3.10	Other <i>a priori</i> Calibrations and Corrections	214
4.4	Solving for Calibration — Bandpass, Gain, Polarization	215
4.4.1	Common Calibration Solver Parameters	215
4.4.1.1	Parameters for Specification : <code>vis</code> and <code>caltable</code>	215
4.4.1.2	Selection: <code>field</code> , <code>spw</code> , <code>selectdata</code> , <code>intent</code> , and <code>observation</code>	215
4.4.1.3	Prior Calibration and Correction: <code>parang</code>	216
4.4.1.4	Previous Calibration: <code>gaintable</code> , <code>gainfield</code> , <code>interp</code> and <code>spwmap</code>	216
4.4.1.5	Solving: <code>solint</code> , <code>combine</code> , <code>preavg</code> , <code>refant</code> , <code>minblperant</code> , <code>minsnr</code>	219
4.4.1.6	Action: <code>append</code> and <code>solnorm</code>	220
4.4.2	Spectral Bandpass Calibration (<code>bandpass</code>)	221
4.4.2.1	Bandpass Normalization	222
4.4.2.2	B solutions	223
4.4.2.3	BPOLY solutions	224
4.4.2.4	What if the bandpass calibrator has a significant slope?	225
4.4.3	Complex Gain Calibration (<code>gaincal</code>)	225
4.4.3.1	Polarization-dependent Gain (G)	226
4.4.3.2	Polarization-independent Gain (T)	228
4.4.3.3	GSPLINE solutions	228
4.4.3.4	Antenna Delays — 'K' solutions	229
4.4.3.5	Cross-Hand Delays — 'KCROSS' solutions	229
4.4.4	Establishing the Flux Density Scale (<code>fluxscale</code>)	230
4.4.4.1	Using Resolved Calibrators	232
4.4.5	Instrumental Polarization Calibration (D,X)	233
4.4.5.1	Heuristics and Strategies for Polarization Calibration	234
4.4.5.2	A Note on channelized polarization calibration	235
4.4.5.3	A Polarization Calibration Example - Circular Feed Basis (e.g., VLA $\nu > 1$ GHz)	236
4.4.5.4	A Polarization Calibration Example - Linear Feed Basis (e.g., ALMA, VLA $\nu < 1$ GHz)	237
4.4.6	Baseline-based Calibration (<code>blcal</code>)	240
4.5	Plotting and Manipulating Calibration Tables	241
4.5.1	Plotting Calibration Solutions (<code>plotcal</code>)	241

4.5.1.1	Examples for <code>plotcal</code>	243
4.5.2	Listing calibration solutions with (<code>listcal</code>)	246
4.5.3	Calibration table statistics (<code>calstat</code>)	248
4.5.4	Calibration Smoothing (<code>smoothcal</code>)	249
4.5.5	Calibration Interpolation and Accumulation (<code>accum</code>)	250
4.5.5.1	Interpolation using (<code>accum</code>)	252
4.5.5.2	Incremental Calibration using (<code>accum</code>)	252
4.6	Application of Calibration to the Data	255
4.6.1	Application of Calibration (<code>applycal</code>)	255
4.6.2	Examine the Calibrated Data	259
4.6.3	Resetting the Calibration Models (<code>delmod</code> and <code>clearcal</code>)	260
4.7	Other Calibration and UV-Plane Analysis Options	261
4.7.1	Splitting out Calibrated uv data (<code>split</code>)	261
4.7.1.1	Averaging in <code>split</code>	262
4.7.2	Recalculation of uvw values (<code>fixvis</code>)	263
4.7.3	Hanning smoothing of uv data (<code>hanningsmooth</code>)	263
4.7.4	MSTransform (<code>mstransform</code>)	264
4.7.4.1	Data selection and re-indexing	265
4.7.4.2	Data partition	266
4.7.4.3	On-the-fly calibration	266
4.7.4.4	Time average	267
4.7.4.5	Channel average	267
4.7.4.6	Hanning smooth	268
4.7.4.7	Combination of spectral windows	268
4.7.4.8	Spectral regridding and reference frame transformation	268
4.7.4.9	Separation of spectral windows	269
4.7.5	Model subtraction from uv data (<code>uvsub</code>)	269
4.7.6	UV-Plane Continuum Subtraction (<code>uvcontsub</code>)	270
4.7.7	Spectral regridding of the MS (<code>cvel</code>)	272
4.7.8	UV-Plane Model Fitting (<code>uvmodelfit</code>)	274
4.7.9	Reweighting visibilities based on their scatter (<code>statwt</code>)	277
4.7.10	Change the signs of visibility phases (<code>conjugatevis</code>)	278
4.7.11	Manipulation of Ephemeris Objects	279
4.7.11.1	Ephemeris tables	279
4.7.11.2	Using <code>fixplanets</code> to attach ephemerides to a field of a Measurement Set	280
4.7.11.3	Use of the ephemeris after attachment	280
4.7.11.4	Spectral frame transformation to the rest frame of the ephemeris object in task <code>cvel</code>	280
4.7.11.5	Ephemerides in ALMA datasets	281
4.8	Examples of Calibration	281

5	Synthesis Imaging	282
5.1	Imaging Tasks Overview	282
5.2	Common Imaging Task Parameters	283
5.2.1	Parameter <code>cell</code>	283
5.2.2	Parameter <code>field</code>	284
5.2.3	Parameter <code>imagename</code>	284
5.2.4	Parameter <code>imsize</code>	284
5.2.5	Parameter <code>mode</code>	285
5.2.5.1	Mode <code>mfs</code>	285
5.2.5.2	Mode <code>channel</code>	287
5.2.5.3	Mode <code>frequency</code>	288
5.2.5.4	Mode <code>velocity</code>	289
5.2.5.5	Sub-parameter <code>interpolation</code>	290
5.2.6	Parameter <code>resmooth</code>	290
5.2.7	Parameter <code>phasecenter</code>	290
5.2.8	Parameter <code>restfreq</code>	291
5.2.9	Parameter <code>spw</code>	291
5.2.10	Parameter <code>stokes</code>	292
5.2.11	Parameter <code>uvtaper</code>	292
5.2.12	Parameter <code>weighting</code>	293
5.2.12.1	'natural' weighting	293
5.2.12.2	'uniform' weighting	294
5.2.12.3	'superuniform' weighting	294
5.2.12.4	'radial' weighting	294
5.2.12.5	'briggs' weighting	295
5.2.12.6	'briggsabs' weighting	295
5.2.13	Parameter <code>vis</code>	296
5.2.14	Primary beams in imaging	296
5.3	Deconvolution using CLEAN (<code>clean</code>)	297
5.3.1	Parameter <code>psfmode</code>	300
5.3.1.1	The <code>clark</code> algorithm	300
5.3.1.2	The <code>hogbom</code> algorithm	301
5.3.1.3	The <code>clarkstokes</code> algorithm	301
5.3.2	The <code>multiscale</code> parameter	301
5.3.3	Parameter <code>gain</code>	302
5.3.4	Parameter <code>imagermode</code>	302
5.3.4.1	Sub-parameter <code>cyclefactor</code>	304
5.3.4.2	Sub-parameter <code>cyclespeedup</code>	305
5.3.4.3	Sub-parameter <code>ftmachine</code>	305
5.3.4.4	Sub-parameter <code>mosweight</code>	306
5.3.4.5	Sub-parameter <code>scaletype</code>	306
5.3.4.6	The <code>threshold</code> revisited	306
5.3.5	Parameter <code>interactive</code>	307
5.3.6	Parameter <code>mask</code>	307
5.3.6.1	Setting <code>clean</code> boxes	308

5.3.6.2	Using clean mask images	308
5.3.7	Parameter <code>minpb</code>	308
5.3.8	Parameter <code>modelimage</code>	308
5.3.9	Parameter <code>niter</code>	309
5.3.10	Parameter <code>pbcor</code>	309
5.3.11	Parameter <code>restoringbeam</code>	309
5.3.12	Parameter <code>threshold</code>	310
5.3.13	Parameter <code>gridmode</code>	310
5.3.14	Interactive Cleaning — Example	311
5.3.15	Mosaic imaging	314
5.3.16	Heterogeneous imaging	317
5.3.17	Polarization imaging	317
5.3.18	Wide-field imaging and deconvolution in <code>clean</code>	317
5.3.18.1	Outlier fields	318
5.3.18.2	Setting up w-projection	320
5.3.18.3	Setting up faceting	320
5.3.18.4	Combination of w-projection and faceting	320
5.4	Refactored <code>clean</code> : <code>tclean</code>	321
5.5	Primary Beam Correction (<code>impbcor</code> , <code>widebandpbcor</code>)	321
5.6	Combined Single Dish and Interferometric Imaging (<code>feather</code>)	322
5.6.1	Visual Interface for <code>feather</code> (<code>casafeather</code>)	324
5.7	Making Deconvolution Masks or Box Regions	325
5.7.1	Making Deconvolution Regions from an Image (<code>boxit</code>)	327
5.8	Insert an Image Model (<code>ft</code>)	327
5.9	Image-plane deconvolution (<code>deconvolve</code>)	328
5.10	Self-Calibration	329
5.11	Parallelized Cleaning	330
5.12	Examples of Imaging	330
6	Image Analysis	331
6.1	Common Image Analysis Task Parameters	333
6.1.1	Input Image (<code>imagename</code>)	333
6.1.2	Region Selection (<code>box</code>)	333
6.1.3	Plane Selection (<code>chans</code> , <code>stokes</code>)	334
6.1.4	Lattice Expressions (<code>expr</code>)	335
6.1.5	Masks (<code>mask</code>)	335
6.1.6	Regions (<code>region</code>)	336
6.2	Image Header Manipulation (<code>imhead</code>)	337
6.2.1	Examples for <code>imhead</code>	338
6.3	Extracting sub-images (<code>imsubimage</code>)	340
6.4	Continuum Subtraction on an Image Cube (<code>imcontsub</code>)	341
6.4.1	Examples for <code>imcontsub</code>	341
6.5	Image-plane Component Fitting (<code>imfit</code>)	342
6.5.1	Examples for <code>imfit</code>	343
6.6	Mathematical Operations on an Image (<code>immath</code>)	345

6.6.1	Examples for <code>immath</code>	346
6.6.1.1	Simple math	346
6.6.1.2	Polarization manipulation	348
6.6.2	Using masks in <code>immath</code>	349
6.7	Computing the Moments of an Image Cube (<code>immoments</code>)	351
6.7.1	Hints for using (<code>immoments</code>)	353
6.7.2	Examples using (<code>immoments</code>)	353
6.8	Generating Position-Velocity Diagrams (<code>impv</code>)	354
6.9	Computing image statistics (<code>imstat</code>)	355
6.9.1	Using the task return value	356
6.9.2	Examples for <code>imstat</code>	359
6.10	Extracting data from an image (<code>imval</code>)	359
6.11	Reordering the Axes of an Image Cube (<code>imtrans</code>)	361
6.12	Collapsing an Image Along an Axis (<code>imcollapse</code>)	363
6.13	Regridding an Image (<code>imregrid</code>)	363
6.14	Redefining the Velocity System of an Image (<code>imreframe</code>)	364
6.15	Rebin an Image (<code>imrebin</code>)	365
6.16	1-dimensional Smoothing (<code>specsmooth</code>)	366
6.17	2-dimensional Smoothing; Image Convolution (<code>imsmooth</code>)	366
6.18	Spectral Line fitting with <code>specfit</code>	368
6.18.1	Polynomial Fits	370
6.18.2	Lorentzian and Gaussian Fits	370
6.18.2.1	One or more single Gaussian/Lorentzian	370
6.18.2.2	Gaussian Multiplets	372
6.18.3	Pixel-by-pixel fits	372
6.19	Calculation of Rotation Measures <code>rmfit</code>	372
6.20	Calculation of Spectral Indices and Higher Order Polynomials (<code>spxfit</code>)	374
6.21	Image Mask Handling <code>makemask</code>	376
6.22	Search for Spectral Line Rest Frequencies (<code>slsearch</code>)	377
6.23	Convert Exported Splatatalogue Catalogs to CASA Tables (<code>splattotable</code>)	379
6.24	Image Import/Export to FITS	379
6.24.1	FITS Image Export (<code>exportfits</code>)	380
6.24.2	FITS Image Import (<code>importfits</code>)	380
6.25	Using the CASA Toolkit for Image Analysis	381
6.26	Examples of CASA Image Analysis	383
7	Visualization With The CASA Viewer	384
7.1	Starting the Viewer	384
7.1.1	Running the CASA Viewer outside <code>casapy</code>	385
7.2	The Viewer Display Panel	386
7.2.1	The Main Toolbar	388
7.2.2	The Mouse Toolbar	389
7.2.3	The Display Area	390
7.2.3.1	The Animator Panel	392
7.2.3.2	The Position Tracking Panel	392

7.2.3.3	The Region Manager Panel	393
7.2.4	Saving and Restoring the Display Panel State	393
7.3	The Data Manager Panel — Saving and Loading Data	394
7.3.1	Loading Data	395
7.3.2	Registered vs. Open Datasets	395
7.3.3	Image Manager	396
7.3.4	Saving Data or Regions	396
7.4	Viewing Images	398
7.4.1	Viewing a Raster Map	398
7.4.1.1	Data Display Options — display and hidden axes	399
7.4.1.2	Data Display Options — basic settings	399
7.4.1.3	Graphical Specification of the Intensity Scale	402
7.4.1.4	Data Display Options — other settings	402
7.4.1.5	Viewer Canvas Manager — Panels, Margins, and Backgrounds	402
7.4.2	Viewing a Contour Map	402
7.4.2.1	Overlay Contours on a Raster Map	405
7.4.3	Regions and the Region Manager	405
7.4.3.1	Region Creation, Selection, and Deletion	407
7.4.3.2	Region Positioning	408
7.4.3.3	Region Statistics	409
7.4.3.4	Saving and Loading Regions	410
7.4.3.5	The Region Fit	410
7.4.3.6	The Region Histogram	411
7.4.4	The Spectral Profile Tool	412
7.4.4.1	Spectral Profile Toolbar	413
7.4.4.2	Main Spectral Profile Window	415
7.4.4.3	Spectral-Line Fitting	417
7.4.4.4	Line Overlays	418
7.4.5	The Brightness Profile Tool	420
7.4.6	The Collapse/Moments Tool	420
7.4.7	The Histogram Tool	422
7.4.8	The Two-D Fitting Tool	423
7.4.9	Interactive Position-Velocity Diagram Creation	424
7.5	Viewing Measurement Sets	425
7.5.1	Data Display Options Panel for Measurement Sets	426
7.5.1.1	MS Options — Basic Settings	426
7.5.1.2	MS Options— MS and Visibility Selections	427
7.5.1.3	MS Options — Display Axes	429
7.5.1.4	MS Options — Flagging Options	430
7.5.1.5	MS Options— Advanced	432
7.5.1.6	MS Options — Apply Button	433
7.6	Printing from the Viewer	433
7.7	Image Viewer (<code>imview</code>)	435
7.8	Measurement Viewer (<code>msview</code>)	437

8	Single-dish Data Processing	439
8.1	Current status, new tasks/operations and CASA SD setup	439
8.1.1	Transition from ASAP <i>Scantable</i> format, and using <i>Scantable</i> format	439
8.1.2	Dictionaries	440
8.1.3	Line Formatting	441
8.1.4	New Tasks, Features and functionality in CASA 4.5	441
8.2	SD data-taking in brief	441
8.2.1	Correlator non-linearity	442
8.3	Overview of SD tools and tasks	443
8.4	Overview for Reducing ALMA Single-dish Data	444
8.4.1	Brief Description of SD reduction	444
8.5	Brief Description of functionality for relevant SD tasks	447
8.5.1	Gridding and imaging: <code>sdimaging</code> , <code>sdtipimaging</code> , <code>sdgrid</code> , <code>sdimprocess</code>	448
8.5.2	Dataset mathematics and manipulation: <code>sdcoadd</code> , <code>sdaverage</code> , <code>tsdsmoother</code> , <code>sdmath</code> , <code>sdscale</code> , <code>sdcoadd</code>	449
8.5.3	Fitting: <code>sdfit</code> , <code>tsdfit</code>	451
8.5.4	Dataset output: <code>sdplot</code> , <code>sdlist</code> , <code>sdstat</code>	451
8.5.5	Data input and output: <code>sdsave</code>	451
8.6	Import of NRO data	451
8.7	Using The ASAP Toolkit	452
8.7.1	ASAP utility and tool descriptions	453
8.7.1.1	<code>Rasterutil</code>	453
8.7.1.2	Data Selection	454
8.7.1.3	State Information	454
8.7.1.4	Masks	456
8.7.1.5	<i>Scantable</i> Management	456
8.7.1.6	<i>Scantable</i> Mathematics	456
8.7.1.7	<i>Scantable</i> Save and Export	457
8.7.2	Calibration	457
8.7.2.1	T_{sys} scaling	457
8.7.2.2	Flux and Temperature Unit Conversion	457
9	Simulation	459
9.1	Simulating ALMA observations with <code>simobserve</code> and <code>simanalyze</code>	460
9.2	Simulating ALMA observations with <code>simalma</code>	462
9.2.1	Implementation details	463
10	Parallel Processing in CASA	466
10.1	The CASA parallelization scheme	466
10.2	Multi-MS creation	467
10.2.1	Partition	467
10.2.1.1	The <code>createmms</code> parameter	467
10.2.2	<code>Importasdm</code>	468
10.3	Parallelization control	468

10.3.1	Requirements	468
10.3.2	Configuration and Start-Up	468
10.3.3	Examples of running CASA in parallel	470
A	Appendix: Obtaining, Installing, and Customizing CASA	472
A.1	Installation On Linux	472
A.1.1	Installation	472
A.1.1.1	Using more than one Linux version of CASA	473
A.1.2	Unsupported platforms	473
A.1.3	Download & Unpack	473
A.2	Installation on Mac OS	474
A.2.0.1	Using more than one Mac version of CASA	475
A.3	Startup	475
A.4	Startup Customization	476
A.5	Updating the data repository	478
B	Appendix: Python and CASA	479
B.1	Python Packages	479
B.2	Automatic parentheses	480
B.3	Indentation	480
B.4	Lists and Ranges	480
B.5	Dictionaries	481
B.5.1	Saving and Reading Dictionaries	482
B.6	Control Flow: Conditionals, Loops, and Exceptions	484
B.6.1	Conditionals	484
B.6.2	Loops	485
B.7	System shell access	486
B.7.1	Using the <code>os.system</code> methods	487
B.7.2	Directory Navigation	488
B.7.3	Shell Command and Capture	489
B.8	Logging	490
B.9	History and Searching	490
B.10	Macros	492
B.11	On-line editing	493
B.12	Executing Python scripts	493
B.13	How do I exit from CASA?	494
C	Appendix: Models, Conventions and Reference Frames	495
C.1	Flux Density Models for setjy	495
C.1.1	Long wavelength calibration	495
C.1.1.1	Baars	496
C.1.1.2	Perley 90	496
C.1.1.3	Perley-Taylor 95	496
C.1.1.4	Perley-Taylor 99	496
C.1.1.5	Perley-Butler 2010	497

C.1.1.6	Perley-Butler 2013	497
C.1.1.7	Scaife-Heald 2012	497
C.1.2	Short wavelength calibration	497
C.2	Velocity Reference Frames	497
C.2.1	Doppler Types	497
C.3	Time Reference Frames	498
C.4	Coordinate Frames	498
C.5	Physical Units	499
C.6	Physical Constants	499
D	Appendix: CASA Region File Format	508
D.1	Region definitions	508
D.2	Allowed shapes	509
D.3	Annotation definitions	510
D.4	Global definitions	511
D.5	Allowed additional parameters	513
D.6	Examples	513
D.7	Fonts and Symbols	514
D.7.1	Allowed symbols	514
D.7.2	Allowed fonts	515
D.7.2.1	Allowed fonts for Linux	515
D.7.2.2	Allowed fonts for MacOS X	515
E	Appendix: The Measurement Equation and Calibration	517
E.1	The HBS Measurement Equation	517
E.2	General Calibrator Mechanics	521
F	Appendix: Data Weights	522
F.1	Introduction	522
F.2	SIGMA and WEIGHT columns	522
F.2.1	Weights in CASA 4.2.1 and Earlier	523
F.2.2	Weights in CASA 4.2.2	524
F.2.3	Weights in CASA 4.3	524
F.2.4	Weights in CASA 4.4 and later	524
G	Appendix: Cal Library Syntax	525
G.1	Introduction	525
G.2	Basic Cal Library Usage	525
G.3	More Advanced Cal Library Usage	527
G.4	General Rules (current, as of CASA 4.5)	527
G.5	Limitations (current, as of CASA 4.5)	528
G.6	Conversion from Existing applycal Scripts	528
G.7	Glossary	529
G.7.1	MS selection	530
G.7.2	Interpolation/application	530
G.7.3	Calibration Mapping	530

H	Appendix: Annotated Example Scripts	531
I	Appendix: CASA Dictionaries	532
I.1	AIPS – CASA dictionary	532
I.2	MIRIAD – CASA dictionary	532
I.3	CLIC – CASA dictionary	532
J	Appendix: Writing Tasks in CASA	535
J.1	The XML file	536
J.2	The <code>task_yourtask.py</code> file	539
J.3	Example: The <code>clean</code> task	539
J.3.1	File <code>clean.xml</code>	540
J.3.2	File <code>task_clean.py</code>	557

List of Tables

2.1	Common MS Columns	87
2.2	Commonly accessed MAIN Table columns	87
4.1	Recognized Flux Density Calibrators. Note that the VLA uses J2000 calibrator names. CASA accepts all strings that contain the names below. E.g. 'PKS 1934-638' will be recognized	200
4.2	'Butler-JPL-Horizons 2012' recognized Solar System Objects for Flux Calibration	201
5.1	Combinations of observing band (wavelength,) and antenna array configurations that require w-projection.	318
C.1	Recognized Flux Density Calibrators. Note that the VLA uses J2000 calibrator names. CASA accepts all strings that contain the names below. E.g. 'PKS 1934-638' will be recognized	496
C.2	Velocity frames in CASA	498
C.3	Doppler types in CASA	498
C.4	Time reference frames in CASA	499
C.5	Coordinate frames in CASA	500
C.6	Prefixes	501
C.7	SI Units	502
C.7	SI Units – continued	503
C.8	Custom Units	504
C.8	Custom Units – continued	505
C.8	Custom Units – continued	506
C.9	Physical Constants	507
F.1	Antenna-based WEIGHT calibration factor definitions for different CASA versions. For System Temperature and Bandpass, k is the channel index. ALMA has channelized T_{sys} ; EVLA does not.	523
I.1	AIPS – CASA dictionary	533
I.2	MIRIAD – CASA dictionary	534
I.3	CLIC–CASA dictionary	534

List of Figures

1.1	Screen shot of the default CASA inputs for task <code>clean</code>	59
1.2	The <code>clean</code> inputs after setting values away from their defaults (blue text). Note that some of the boldface ones have opened up new dependent sub-parameters (indented and green).	60
1.3	The <code>clean</code> inputs where one parameter has been set to an invalid value. This is drawn in red to draw attention to the problem. This hapless user probably confused the 'hogbom' clean algorithm with Harry Potter.	61
1.4	The CASA Logger GUI window under Linux. Note that under MacOSX a stripped down logger will instead appear as a Console.	66
1.5	Using the Search facility in the <code>casalogger</code> . Here we have specified the string 'apply' and it has highlighted all instances in green.	67
1.6	Using the <code>casalogger</code> Filter facility. The log output can be sorted by Priority, Time, Origin, and Message. In this example we are filtering by Origin using 'clean', and it now shows all the log output from the <code>clean</code> task.	68
1.7	CASA Logger - Insert facility: The log output can be augmented by adding notes or comments during the reduction. The file should then be saved to disk to retain these changes.	69
1.8	Flow chart of the data processing operations that a general user will carry out in an end-to-end CASA reduction session.	73
2.1	The contents of a Measurement Set. These tables compose a Measurement Set named <code>ngc5921.demo.ms</code> on disk. This display is obtained by using the File:Open menu in <code>browsetable</code> and left double-clicking on the <code>ngc5921.demo.ms</code> directory.	86
3.1	A freshly-started <code>plotms</code> GUI window. Note that the Plots > Data tab is selected, which is discussed in § 3.3.1.1, 3.3.1.7, and 3.3.1.9.	126
3.2	The Plots > Axes tab in the <code>plotms</code> GUI window, used to make a plot of Amp versus Channel	129
3.3	Plot of amplitude versus time, before (left) and after (right) flagging two marked regions. To unflag these regions, mark the two same regions and click the Unflag button.	134
3.4	Overplotting in <code>plotms</code> : Two different y-axes have been chosen for this plot, amplitude and elevation.	141

3.5	The <code>plotxy</code> plotter, showing the Jupiter data versus uv-distance. You can see bad data in this plot. The bottom set of buttons on the lower left are: 1,2,3) Home, Back, and Forward . Click to navigate between previously defined views (akin to web navigation). 4) Pan . Click and drag to pan to a new position. 5) Zoom . Click to define a rectangular region for zooming. 6) Subplot Configuration . Click to configure the parameters of the subplot and spaces for the figures. 7) Save . Click to launch a file save dialog box. The upper set of buttons in the lower left are: 1) Mark Region . Press this to begin marking regions (rather than zooming or panning). 2,3,4) Flag, Unflag, Locate . Click on these to flag, unflag, or list the data within the marked regions. 5) Next . Click to move to the next in a series of iterated plots. Finally, the cursor readout is on the bottom right.	144
3.6	The <code>plotxy</code> iteration plot. The first set of plots from the example in § 3.3.2.3 with <code>iteration='antenna'</code> . Each time you press the Next button, you get the next series of plots.	149
3.7	Multi-panel display of visibility versus channel (top), antenna array configuration (bottom left) and the resulting uv coverage (bottom right). The commands to make these three panels respectively are: 1) <code>plotxy('ngc5921.ms', xaxis='channel', datacolumn='data', field='0', subplot=211, plotcolor="", plotsymbol='go')</code> 2) <code>plotxy('ngc5921.ms', xaxis='x', field='0', subplot=223, plotsymbol='r.')</code> 3) <code>plotxy('ngc5921.ms', xaxis='u', yaxis='v', field='0', subplot=224, plotsymbol='b', figfile='ngc5921_multiplot.png')</code>	152
3.8	Plot of amplitude versus uv distance, before (left) and after (right) flagging two marked regions. The call was: <code>plotxy(vis='ngc5921.ms', xaxis='uvdist', field='1445*')</code>	154
3.9	This figure shows the geometry used to compute shadowed antennas.	166
3.10	This screenshot represents a run where <code>'tfcrop'</code> was run on a <code>spw='9'</code> with mainly narrow-band RFI. RIGHT : An example of protecting a spectral line (in this case, demonstrated on an RFI spike) by setting the <code>spw-selection</code> to <code>spw='0:0 45;53 63'</code> . In both figures, the top row indicates the data before flagging, and the bottom row after flagging.	169
3.11	Example of <code>rflag</code> on narrow-band RFI	171
3.12	This screenshot represents a run where <code>'tfcrop'</code> was run only on <code>'ABS_RR'</code> (top row) and followed by an extension along time and correlations (bottom row).	173
3.13	<code>browsetable</code> : The browser displays the main table within a frame. You can scroll through the data (x=columns of the <code>MAIN</code> table, and y=the rows) or select a specific page or row as desired. By default, 1000 rows of the table are loaded at a time, but you can step through the MS in batches.	183
3.14	<code>browsetable</code> : You can use the tab for <code>Table Keywords</code> to look at other tables within an MS. You can then double-click on a table to view its contents.	184
3.15	<code>browsetable</code> : Viewing the <code>SOURCE</code> table of the MS.	185
4.1	Flow chart of synthesis calibration operations. Not shown are use of table manipulation and plotting tasks <code>accum</code> , <code>plotcal</code> , and <code>smoothcal</code> (see Figure 4.2).	189

4.2	Chart of the table flow during calibration. The parameter names for input or output of the tasks are shown on the connectors. Note that from the output solver through the accumulator only a single calibration type (e.g. 'B', 'G') can be smoothed, interpolated or accumulated at a time. <code>accum</code> is optional (and not recommended as of v4.0). The final set of cumulative calibration tables of all types (accumulated or as a list of <code>caltables</code>) are then input to <code>applycal</code> as shown in Figure 4.1.	192
4.3	The weather information for a MS as plotted by the task <code>plotweather</code>	199
4.4	Display of the amplitude (upper) and phase (lower) gain solutions for all antennas and polarizations in the <code>ngc5921</code> <code>post-fluxscale</code> table.	244
4.5	Display of the amplitude (upper), phase (middle), and signal-to-noise ratio (lower) of the <code>bandpass</code> 'B' solutions for <code>antenna='0'</code> and both polarizations for <code>ngc5921</code> . Note the falloff of the SNR at the band edges in the lower panel.	245
4.6	Display of the amplitude of the <code>bandpass</code> 'B' solutions. Iteration over antennas was turned on using <code>iteration='antenna'</code> . The first page is shown. The user would use the Next button to advance to the next set of antennas.	246
4.7	The ' <code>amp</code> ' of gain solutions for <code>NGC4826</code> before (top) and after (bottom) smoothing with a 7200 sec <code>smoothtime</code> and <code>smoothtype='mean'</code> . Note that the first solution is in a different <code>spw</code> and on a different source, and is not smoothed together with the subsequent solutions.	250
4.8	The ' <code>phase</code> ' of gain solutions for <code>NGC4826</code> before (top) and after (bottom) ' <code>linear</code> ' interpolation onto a 20 sec <code>accumtime</code> grid. The first scan was <code>3C273</code> in <code>spw='0'</code> while the calibrator scans on <code>1331+305</code> were in <code>spw='1'</code> . The use of <code>spwmap</code> was necessary to transfer the interpolation correctly onto the <code>NGC4826</code> scans.	253
4.9	The final ' <code>amp</code> ' (top) and ' <code>phase</code> ' (bottom) of the self-calibration gain solutions for Jupiter. An initial phase calibration on 10s <code>solint</code> was followed by an incremental gain solution on each scan. These were accumulated into the cumulative solution shown here.	256
4.10	The final ' <code>amp</code> ' versus ' <code>uvdist</code> ' plot of the self-calibrated Jupiter data, as shown in <code>plotxy</code> . The ' <code>RR LL</code> ' correlations are selected. No outliers that need flagging are seen.	260
4.11	Use of <code>plotxy</code> to display corrected data (red and blue points) and uv model fit data (green circles).	277
5.1	Close-up of the top of the interactive <code>clean</code> window. Note the boxes at the left (where the <code>iterations</code> , <code>cycles</code> , and <code>threshold</code> can be changed), the buttons that control add/erase, the application of mask to channels, and whether to stop, complete, or continue cleaning, and the row of Mouse-button tool assignment icons.	303
5.2	Screen-shots of the interactive <code>clean</code> window during deconvolution of the VLA 6m Jupiter dataset. We start from the calibrated data, but before any self-calibration. In the initial stage (left), the window pops up and you can see it dominated by a bright source in the center. Next (right), we zoom in and draw a box around this emission. We have also at this stage dismissed the tape deck and Position Tracking parts of the display (§ 7.2) as they are not used here. We have also changed the <code>iterations</code> to 30 for this boxed clean. We will now hit the Next Action Continue Cleaning button (the green clockwise arrow) to start cleaning.	312

5.3	We continue in our interactive cleaning of Jupiter from where Figure 5.2 left off. In the first (left) panel, we have cleaned 30 iterations in the region previously marked, and are zoomed in again ready to extend the mask to pick up the newly revealed emission. Next (right), we have used the Polygon tool to redraw the mask around the emission, and are ready to Continue Cleaning for another 100 iterations. . . .	313
5.4	We continue in our interactive cleaning of Jupiter from where Figure 5.3 left off. In the first (left) panel, it has cleaned deeper, and we come back and zoom in to see that our current mask is good and we should clean further. We change npercycle to 500 (from 100) in the box at upper right of the window. In the final panel (right), we see the results after this clean. The residuals are such that we should terminate the clean using the red X button and use our model for self-calibration.	314
5.5	After clean and self-calibration using the intensity image, we arrive at the final polarization image of Jupiter. Shown in the viewer superimposed on the intensity raster is the linear polarization intensity (green contours) and linear polarization B-vectors (vectors). The color of the contours and the sampling and rotation by 90 degrees of the vectors was set in the Display Options panel. A LEL expression was used in the Load Data panel to mask the vectors on the polarized intensity.	315
5.6	Screen-shot of the interactive clean window during deconvolution of the NGC5921 spectral line dataset. Note where we have selected the mask to apply to the Displayed Plane rather than All Channels . We have just used the Polygon tool to draw a mask region around the emission in this channel, which will apply to this channel only.	316
5.7	Visual “casafeather” interface to the feather task.	325
5.8	The scatter plot in casafeather.	326
5.9	The casafeather “customize” window.	326
6.1	NGC2403 VLA moment zero (left) and NGC4826 BIMA moment one (right) images as shown in the viewer	354
7.1	The Viewer Display Panel (left) and the Data Manager (right) panel for a regular image or data cube.	384
7.2	The Viewer Display Panel (left) and Data Display Options (right) panels with a Measurement Set open.	386
7.3	The display panel’s Main Toolbar appears directly below the menus and contains ‘shortcut’ buttons for most of the frequently-used menu items.	388
7.4	The ‘Mouse Tool’ Bar allows you to assign how mouse buttons behave in the image display area. Initially, zooming, color adjustment, and rectangular regions are assigned to the left, middle and right mouse buttons. Click on a tool with a mouse button to assign that tool to that mouse button.	389
7.5	”Preferences” dialog to manually change the docking and size of the viewer panel. .	391
7.6	The animator panel, which allows one to scroll along the <i>z</i> axis of a data cube (using the Channels tape deck) or cycle among open Images. The panel can be undocked from the main display panel.	392

7.7	The position tracking panel, which gives information about the open data cube at the current location of the cursor. Freeze the position tracking panel using the SPACE bar.	393
7.8	The load tab of the Data Manager panel. This appears if you open the viewer without any infile specified, if you use select Open from the Data drop down menu, or click the Open (Folder) icon. You can access the save image or save region tabs from this view or by selecting Save as... from the Data drop down menu. The load tab shows all files in the current directory that can be loaded into the viewer — images, MS, CASA region files, and Display Panel State files.	394
7.9	The Image Manager.	397
7.10	The Save Data panel that appears when selecting the 'Save as...' (Figure 7.3).	397
7.11	The basic settings category of the Data Display Options panel and the interactive tool for setting the mapping from intensity to color.	399
7.12	Example curves for scaling power cycles	401
7.13	A multi-panel display set up through the Viewer Canvas Manager	403
7.14	The Viewer Display Panel (left) and Data Display Options panel (right) after choosing contour map from the Load Data panel. The image shown is for channel 11 of the NGC5921 cube, selected using the Animator tape deck, and zoomed in using the tool bar icon. Note the different options in the open basic settings category of the Data Display Options panel (as compared to raster image in Figure 7.1).	404
7.15	The Viewer Display Panel (left) and Data Display Options panel (right) after overlaying a Contour Map of velocity on a Raster Image of intensity. The image shown is for the moments of the NGC5921 cube, zoomed in using the tool bar icon. The tab for the contour plot is open in the Data Display Options panel.	406
7.16	The Region Manager Panel, which becomes active once at least one region is created. Cycle through available regions using the slider bar at the bottom and use the various tabs to adjust, analyze, load, and save regions.	406
7.17	Selecting an image region (done with SHIFT+click). The region can be resized by dragging the handles or deleted by hitting ESCAPE.	407
7.18	The positioning tab in the Region Manager. Use it to manually adjust the location, width, and display style of the selected region.	408
7.19	The statistics tab in the Region Manager.	409
7.20	The save/load tab in the Region Manager.	410
7.21	The histogram tab in the Region Manager. Right click to zoom. Hit SHIFT + Right Click to adjust the details of the histogram display.	411
7.22	The Spectral Profile panel (right) that appears when pressing the button Open the Spectrum Profiler in the Main Toolbar and then use the tools to select a region in the image, such as the rectangular region on the left panel. The Spectral Profile tool shows the spectrum of the most recent region highlighted and updates to track movements of the region if moved by dragging with the mouse.	412
7.23	The toolbar for the Spectral Profile tool allows the user to save the spectrum, print or save the tool as an image, edit preferences (general, tool, legend), spectral smoothing, pan or zoom around the spectrum, select a range of interest, jump to a channel, or add a label.	413

7.24	Preferences options in the Spectral Profile Tool. From the toolbar, one can access dialogs to set overall viewer preferences, colors for plotting, how the plot legend is displayed, and spectral smoothing method and kernel width.	414
7.25	The main panel for the Spectral Profile tool. Buttons along the bottom row allow the axes to be set. Arrow keys pan and dragging out an area with the mouse zooms. Holding CTRL and right clicking in the spectrum will jump the main Viewer Display panel to display that frequency channel.	416
7.26	The Spectral Line Fitting tab in the Spectral Profile Tool. The user can fit a combination of a polynomial and multiple Gaussian components, specifying the range to be fit (gray region) manually or with a shift+click+drag. Initial estimates for each component may be entered by hand or specified via an initial estimates GUI. The results are output to a dialog and text file with the fit overplotted (here in blue) on the spectrum (with the possibility to save it to disk).	417
7.27	The left panel shows the graphical specification of initial estimates for Gaussian fitting. Slider bars specify the center, FWHM, and peak intensity for the initial estimate. The right panel shows the verbose output of the fitting.	418
7.28	Line Overlays in the Spectral Profile Tool. The Line Overlay tab, shown at the bottom, allows users to query the CASA copy of the Spaltalogue spectral line database. Enter the redshift of your source (right panel), select and Astronomical Filter from the drop down menu, and use shift+click+drag to select a frequency range (or do so manually). The "Search" button will bring up the dialog seen at the left top part of the image, which can in turn be used to graph the candidate lines in the main Spectral Profile window (here CO v=0).	419
7.29	1-dimensional slice of an image. The 1D slicer tool shows the brightness distribution along line segments.	420
7.30	The Collapse/Moments tool, accessed from the Main Toolbar or the Tools drop down menu. The mean spectrum from the region in the Main Display Panel appears in the top part of the tool. After selecting a range, a moment to calculate, and optionally data to exclude click collapse to calculate a new image.	421
7.31	The Histogram tool, accessed from the Main Toolbar or the Tools drop down menu. Details of the display and included pixels can be manipulated via the menus along the top of the window. The right hand panel allows one to attempt to fit a distribution to the histogram.	422
7.32	The interface to the two dimensional fitting tool (Tools:Fit... or the blue circles icon). The interface allows you to specify and automatically generate (Find Sources) initial estimates, to specify the range of pixel values to be included in the fit, and to specify the output (log file, residual image, and visualization). Click Fit to start the fit.	423
7.33	Interactive creation of position-velocity cuts in the viewer. Use the P/V tool from the Mouse Toolbar to define a cut, then use the pV tool from the Region Manager Panel to adjust the cut (including the width). Click Generate P/V to build the position velocity cut and open it in a new Viewer Display Panel (from which it can be saved to disk).	424

7.34	The Load Data - Viewer panel as it appears if you select an MS. The only option available is to load this as a Raster Image . In this example, clicking on the Raster Image button would bring up the displays shown in Figure 7.2.	425
7.35	The MS for NGC4826 BIMA observations has been loaded into the viewer. We see the first of the spw in the Display Panel, and have opened up MS and Visibility Selections in the Data Display Options panel. The display panel raster is not full of visibilities because spw 0 is continuum and was only observed for the first few scans. This is a case where the different spectral windows have different numbers of channels also.	428
7.36	The MS for NGC4826 from Figure 7.35, now with the Display Axes open in the Data Display Options panel. By default, channels are on the Animation Axis and thus in the tapedeck, while spectral window and polarization are on the Display Axes sliders.	429
7.37	The MS for NGC4826, continuing from Figure 7.36. We have now put spectral window on the Animation Axis and used the tapedeck to step to spw 2 , where we see the data from the rest of the scans. Now channels is on a Display Axes slider, which has been dragged to show Channel 33	430
7.38	Printing the display to a hardcopy of a file. From the Viewer Print Manager , located in top right here and accessed by the print icon or from the Data drop down menu, you can use the Save button to save an image or Print directly to a printer. To achieve the best results, it is often helpful to adjust the settings in the Data Display Options and Viewer Canvas Manager , shown at right.	434
7.39	Data selection in msview	438
8.1	Observations modes for science observations <i>A) Left</i> where the OFF position is distant from the target field, and calibrator observations <i>B), Right</i> where the OFF positions are extracted from the edges of the scanned area.	442
8.2	General steps for processing both <i>calibrator</i> and <i>science</i> data, and applying calibrations to science data. The right-most column describes the CASA 4.5 tasks called during each stage.	445

Chapter 1

Introduction

This document describes how to calibrate and image interferometric and single-dish radio astronomical data using the CASA (Common Astronomy Software Application) package. CASA is a suite of astronomical data reduction tools and tasks that can be run via the IPython interface to Python. CASA is being developed in order to fulfill the data post-processing requirements of the ALMA and EVLA projects, but also provides basic and advanced capabilities useful for the analysis of data from other radio, millimeter, and submillimeter telescopes.

You have in your hands the latest release of CASA. This package is under active development, and thus there are a number of caveats and limitations for the use of this package. See the release notes (§ 1.1) below for more information, and pay heed to the numerous **ALERT**s placed throughout this reference. You can expect regular updates and patches, as well as increasing functionality.

This user reference and cookbook is a task-based walk-through of interferometric data reduction and analysis. In CASA, **tasks** represent the more streamlined operations that a typical user would carry out. The idea for having tasks is that they are simple to use, provide a more familiar interface, and are easy to learn for most astronomers who are familiar with radio interferometric data reduction (and hopefully for novice users as well). In CASA, the **tools** provide the full capability of the package, and are the atomic functions that form the basis of data reduction. These tools augment the tasks, or fill in gaps left by tasks that are under development but not yet available. See the **CASA Toolkit Manual** for more details on the tools (available from casa.nrao.edu). Note that in most cases, the tasks are Python interface scripts to the tools, but with specific, limited access to them and a standardized interface for parameter setting. The tasks and tools can be used together to carry out more advanced data reduction operations.

Inside the Toolkit:

Throughout this user reference, we will occasionally intersperse boxed-off pointers to parts of the toolkit that power users might want to explore.

For the moment, the audience is assumed to have some basic grasp of the fundamentals of synthesis imaging, so details of how a radio interferometer or telescope works and why the data needs to undergo calibration in order to make synthesis images are left to other documentation — a good

place to start might be *Synthesis Imaging in Radio Astronomy II* (1999, ASP Conference Series Vol. 180, eds. Taylor, Carilli & Perley).

This reference is broken down by the main phases of data analysis:

- data import, export, and selection (Chapter 2),
- examination and flagging of data (Chapter 3),
- interferometric calibration (Chapter 4),
- interferometric imaging (Chapter 5),
- image analysis (Chapter 6), and
- data and image visualization (Chapter 7).
- single dish data analysis (Chapter 8), and
- simulation (Chapter 9).

These are included for users that will be doing EVLA and ALMA telescope commissioning and software development.

The general appendices provide more details on what's happening under the hood of CASA, as well as supplementary material on tasks, scripts, and relating CASA to other packages. These appendices include:

- obtaining and installing CASA (Appendix A),
- more details about Python and CASA (Appendix B),
- a list of used models, conventions, and reference frames (Appendix C),
- a description of the CASA region format (Appendix D),
- a discussion of the Hamaker-Bregman-Sault Measurement Equation (Appendix E),
- annotated scripts for typical data reduction cases (Appendix H), and
- CASA dictionaries to AIPS, MIRIAD, and CLIC (Appendix I).
- Writing your own CASA Task (Appendix J).

The CASA User Documentation includes:

- **CASA User Reference & Cookbook** — this document, a task-based data analysis walk-through and instructions;
- **CASA in-line help** — accessed using `help` in the `casapy` interface;

- The **CASA Toolkit Reference Manual** — details on a specific task or tool does and how to use it.
- The **CASA Task Reference Manual** — the information from the inline help and task documentation, available online in HTML.

The CASA home page can be found at:

`http://casa.nrao.edu`

From there you can find documentation and assistance for the use of the package, including the User Documentation. You will also find information on how to obtain the latest release and receive user support.

There is also a CASAGuides Wiki

`http://casaguides.nrao.edu`

that contains helpful information on CASA startup, AIPS-to-CASA cheat sheet, example scripts of processing your data in CASA, along with hints and tricks to best use this package.

1.0.1 Reference for Publications

If you use CASA for any of your data reduction or analysis, you may use the following reference:

McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, Astronomical Data Analysis Software and Systems XVI (ASP Conf. Ser. 376), ed. R. A. Shaw, F. Hill, & D. J. Bell (San Francisco, CA: ASP), 127

1.1 About This Release

The 4.5.2 CASA version is a second patch to the 4.5.0 CASA release, which is described below, together with the CASA 4.5.0 release notes.

We occasionally issue patches and 'monthly' versions of CASA. To get notified, please subscribe to the 'casa-users' mailing list. 'Monthlies' as well as CASA releases are available at `http://casa.nrao.edu`. Releases will be announced via the 'casa-announce' mailing list. To subscribe, please visit `http://casa.nrao.edu`. For feedback, and help please go to the NRAO helpdesk `http://help.nrao.edu`; for ALMA questions please use the ALMA helpdesk `http://help.almascience.org`.

ALERT

Boxes like this will bring to your attention some of the features (or lack thereof) in the current release of CASA. There are also **ALERT** notes in the text.

Note that in its current incarnation CASA is designed to support Karl G. Jansky VLA, ALMA, and older VLA data, as well as single dish data from ALMA as well as Nobeyama. Data from other telescopes, be it single dish or interferometers can be imported from `uvfits`, `FITS-IDI`, or `sdfits` formats into Measurements Sets (MS) or scantables in CASA. Given the variety of non-standard fits formats, we cannot guarantee that CASA will fully support data from all telescopes. However, efforts are made to support data formats from other facilities.

1.2 Obtaining CASA

CASA is available for the following operating systems:

- Linux
 - RedHat 5 and 6 (64-bit)
- Mac OS
 - Mac OS 10.9/10.10 (Maverick/Yosemite)

The above OSs are the ones that we use to test the CASA package. Other flavors of Linux may work, too. We do not support Mac OS 10.9 or 10.10 at this stage.

The latest and previous releases can be downloaded from our CASA home page: <http://casa.nrao.edu>, following the 'Obtaining CASA' link (direct link: http://casa.nrao.edu/casa_obtaining.shtml).

1.2.1 What's New in Release 4.5.2

CASA 4.5.2 patch: CASA 4.5.2 is a second patch building upon CASA 4.5.1. The additional changes are mostly affecting the ALMA pipeline, in particular to accommodate updated observing strategies. It also fixes a bug in `tclean` to create a primary beam when using `chanchunks`, and addresses some edge cases in `plotbandpass`.

CASA 4.5.1 patch: As mentioned above, the CASA 4.5.1 patch to the CASA 4.5.0 release incorporates a new version of the ALMA data reduction pipeline (see the ALMA pipeline group for details). We fixed a bug in `setjy` that caused positional offsets when transforming solar system ephemeris data from J2000 to ICRS. We also introduced channel chunking and memory management improvements to `tclean`.

The CASA 4.5.0 release emphasized improved stability, robustness, and code refactoring. Many bugs were fixed and development such as parallelization of the calibration steps are being introduced. `tclean` will eventually replace `clean` and has underwent a period of intensive testing and improvement.

Major feature improvements over the previous version of CASA include:

- **Imaging**

- `tclean` is now being released, which is a first version of a new imaging task, eventually replacing `clean`. `tclean` is still undergoing testing and improvement. It has a new, more straightforward interface, allows more combinations of algorithms as well as new options. `tclean` has also been designed to support parallel computing.

- **Data examination/editing/import/export**

- `importasdm` can now create a multi-MS (MMS) used for parallel processing
- Added support for time averaging in `clip` mode inside `flagdata`
- Added code to the `importasdm` task to recalculate the UVW coordinates for all fields with attached ephemerides
- `importfitsidi` has a new parameter `constobsid` to specify if different observation IDs should be written in separate files, as well as `scanreindexgap_s` to specify the time gap between scans
- `plotms` can now be used for read-only files
- `flagdata`'s `summary` mode now also shows a spw per field breakdown
- expanded number of methods for the `msmd` tool for expanded MS metadata capture (including observation, polarization, source, field data and more)

- **Calibration**

- The Calibration Library can now be specified in `plotms` from the command line and the GUI
- Calibration instructions can also be specified directly in the `plotms callib` parameter
- Delay, antpos, opac, and gaincurve caltables now work via the Cal Library
- `setjy` is now using the flux standard 'Perley Butler 2013' as default

- **Simulation**

- seeds are now available in `ia.addnoise`

- **Data and Image Manipulation**

- `mstransform` can now average across fields

- **Data analysis**

- The `chans` and `stokes` selection parameters in image analysis tasks are now overruling any specified selection in the region files
- `imval` is now also working on position-velocity images

- **Single Dish**

- new single dish tasks that operate on MSes are available:

- * a new task, `tsdsmoother`, is available for smoothing of spectra. The available smoothing kernel is `kernel = 'gaussian'`
 - * a new task, `tsdfit`, is available to get fitting coefficients of spectral lines. The supported fitting function is `fitfunc='gauss'` (Gaussian fitting)
 - enhancements to single dish tasks those operate on MSes
 - * a new mode, `maskmode='auto'`, is available in task `tsdbaseline`. The mode supports automatic detection of frequency ranges of baseline using a line finder
 - * a new mode, `blmode='apply'`, is available in task `tsdbaseline` for the application of baseline table generated by `blmode='fit'`
 - * new formats, `bloutput='csv'` (comma separated values) and `'text'` (an ascii text), are available in `tsdbaseline` to store fitting coefficients for each spectrum. Composite of available `bloutput` formats are also supported
 - a new parameter, `intent`, in `sdimaging` to support data selection by scan intent
- **Performance**
 - Parallelized processing of the pre-imaging calibration steps is now available
 - **Other**
 - all GUIs are able to identify their CASA version number
 - **Experimental tasks**
 - `cvel2` to provide the functionality of `cvel` based on the `mstransform` task
 - `split2` to provide the functionality of `split` based on the `mstransform` task
 - `hanningsmooth2` to provide the functionality of `hanningsmooth` based on the `mstransform` task

For known issues with this release please visit the CASA webpages:

http://casa.nrao.edu/release_ki.shtml

1.3 CASA Basics — Information for First-Time Users

This section assumes that CASA has been installed on your LINUX or OSX system. See Appendix A for instructions on how to obtain and install CASA.

1.3.1 Before Starting CASA

First, you will most likely be starting CASA running from a working directory that has your data in it, or at least where you want your output to go. It is easiest to start from there rather than changing directories inside `casapy`. **ALERT:** There is at least one task (`plotxy`) that fails if the

path to your working directory contains spaces in its name, e.g. `/users/smyers/MyTest/` is fine, but `/users/smyers/My Test/` is not! Please use our new task `plotms` whenever possible and we may work on a better handling of spaces in path names.

If you have done a default installation under Linux using rpms, or on the Mac with the CASA application, then there should be a `sh` script called `casapy` in the `/usr/bin` area which is in your path. This shell will set up its environment and run the version of `casapy` that it points to. If this is how you set up the system, then you need to nothing further and can run `casapy`.

For internal NRAO users we keep different version of CASA, the latest “casapy” release, the “casapy-stable” “Stable” version that is more developed than the Release but without the full documentation (and no GUI testing). We also offer the “Test” version, “casapy-test”, which is produced on a roughly weekly basis with all the latest code but it underwent much less rigorous testing. Instructions how to run the different versions at NRAO can be found on our <http://casa.nrao.edu> webpages under the “CASA at NRAO” link for the different NRAO sites.

1.3.1.1 Environment Variables

Before starting up `casapy`, you should set or reset any *environment variables* needed, as CASA will adopt these on startup. For example, the `PAGER` environment variable determines how help is displayed in the CASA terminal window (see § 1.3.8.3). The choices are `less`, `more`, and `cat`.

In `bash`, pick one of

```
PAGER=less
PAGER=more
PAGER=cat
```

followed by

```
export PAGER
```

In `csh` or `tcsh`, pick one of

```
setenv PAGER less
setenv PAGER more
setenv PAGER cat
```

The actions of these are as if you were using the equivalent Unix shell command to view the help material. See § 1.3.8.3 for more information on these choices. We recommend using the `cat` option for most users, as this works smoothly both interactively and in scripts.

Some CASA processes will work on a large number of temporary files. The OS, however, may have a built-in limit on the number of such files. We recommend to increase the limit to > 1024 . A command like

```
ulimit -n 2048
```

should give CASA enough accessible files to run successfully.

1.3.1.2 Where is CASA?

Note that the path to the CASA installation, which contains the scripts and data repository, will also depend upon the installation. With a default installation under Linux this will probably be in

```
/usr/lib64/casapy/
```

If the unpacked tarball is placed somewhere else, one may add the `PATH` variable to include, e.g.

```
export PATH=$PATH:/<path>/casapy-<version>/bin
```

for bash and

```
set path = ($path /<path>/casapy-<version>/bin)
```

in csh shell.

In a Mac OSX default install it will likely be an application in the Applications folder.

You can find the location after initialized by looking at the `CASAPATH` environment variable. You can find it within `casapy` by

```
pathname=os.environ.get('CASAPATH').split()[0]
print pathname
```

1.3.2 Starting CASA

After having run the appropriate `casainit` script, CASA is started by typing

```
casapy
```

on the UNIX command line, e.g.

```
casapy
```

After startup information, you should get an IPython

```
CASA <1>:
```

command prompt in the xterm window where you started CASA. CASA will take approximately 10 seconds to initialize at startup in a new working directory; subsequent startups are faster. CASA is active when you get a

```
CASA <1>
```

prompt in the command line interface. You will also see a `logger` GUI appear on your Desktop (usually near the upper left).

You also have the option of starting CASA with various `logger` options (see § 1.5.2.1). For example, if you are running remotely in a terminal window without an X11 connection, or if you just do not want to see the `logger` GUI, and want the `logger` messages to come to your terminal, do

```
casapy --nologger --log2term
```

See § 1.5.2 for information on the `logger` in general.

1.3.3 Ending CASA

You can exit CASA by typing `quit`. This will bring up the query

```
Do you really want to exit ([y]/n)?
```

to give you a chance in case you did not mean to exit. You can also quit using `%exit` or `CTRL-D`.

If you don't want to see the question "Do you really want to exit [y]/n?", then just type `Exit` or `exit` and CASA will stop right then and there.

1.3.4 What happens if something goes wrong?

ALERT: Please check the CASA Home Page for Release Notes and FAQ information including a list of known problems. If you think you have encountered an unknown problem, please consult the CASA HelpDesk (contact information on the CASA Home Page). See also the caveats to this Release (§ 1.1 for pointers to our policy on User Support).

First, always check that your inputs are correct; use the

```
help <taskname>
```

(§ 1.3.8.2) or

```
help par.<parameter name>
```

(§ 1.3.8.4) to review the inputs/output.

1.3.5 Aborting CASA execution

If something has gone wrong and you want to stop what is executing, then typing `CTRL-C` (Control and C keys simultaneously) will usually cleanly abort the application. This will work if you are running a task synchronously. If this does not work on your system then try `CTRL-Z` to put the task or shell in the background, and then follow up with a `kill -9 <PID>` where you have found the relevant `casapy` process ID (PID) using `ps` (see § 1.3.6 below).

If the problem causes CASA to crash, see the next sub-section.

See § 1.4.2 for more information on running tasks.

Alert: `CTRL-C` while a task runs can corrupt your input data file, e.g. when a scratch column is filled while aborting. If in doubt, wait until the task has finished, delete the new files produced, and start again.

1.3.6 What happens if CASA crashes?

Usually, restarting `casapy` is sufficient to get you going again after a crash takes you out of the Python interface. Note that there may be spawned subprocesses still running, such as the `casaviewer` or the `logger`. These can be dismissed manually in the usual manner. After a crash, there may also be hidden processes. You can find these by listing processes, e.g. in linux:

```
ps -elf | grep casa
```

or on MacOSX (or other BSD Unix):

```
ps -aux | grep casa
```

You can then kill these, for example using the Unix `kill` or `killall` commands. This may be necessary if you are running remotely using `ssh`, as you cannot logout until all your background processes are terminated. For example,

```
killall ipcontroller
```

or

```
killall Python
```

will terminate the most common post-crash zombies.

1.3.7 Python Basics for CASA

Within CASA, you use Python to interact with the system. This does not mean an extensive Python course is necessary - basic interaction with the system (assigning parameters, running tasks) is straightforward. At the same time, the full potential of Python is at the more experienced user's disposal. Some further details about Python, IPython, and the interaction between Python and CASA can be found in Appendix B.

The following are some examples of helpful hints and tricks on making Python work for you in CASA.

1.3.7.1 Variables

Python variables are set using the `<parameter> = <value>` syntax. Python assigns the type dynamically as you set the value, and thus you can easily give it a nonsensical value, e.g.

```
vis = 'ngc5921.ms'  
vis = 1
```

The CASA parameter system will check types when you run a task or tool, or more helpfully when you set inputs using `inp` (see below). CASA will check and protect the assignments of the global parameters in its namespace.

Note that Python variable names are case-sensitive:

```
CASA <109>: Foo = 'bar'
CASA <110>: foo = 'Bar'
CASA <111>: foo
  Out[111]: 'Bar'
CASA <112>: Foo
  Out[112]: 'bar'
```

so be careful.

Also note that misspelling a variable assignment will not be noticed (as long as it is a valid Python variable name) by the interface. For example, if you wish to set `correlation='RR'` but instead type `correlation='RR'` you will find `correlation` unset and a new `correlation` variable set. Command completion (see §1.3.8.1) should help you avoid this.

1.3.7.2 Lists and Ranges

Sometimes, you need to give a task a list of indices. If these are consecutive, you can use the Python `range` function to generate this list:

```
CASA <1>: iflist=range(4,8)
CASA <2>: print iflist
[4, 5, 6, 7]
CASA <3>: iflist=range(4)
CASA <4>: print iflist
[0, 1, 2, 3]
```

See Appendix B.4 for more information.

1.3.7.3 Indexes

As in C, Python indices are 0-based. For example, the first element in a list `antlist` would be `antlist[0]`:

```
CASA <113>: antlist=range(5)
CASA <114>: antlist
  Out[114]: [0, 1, 2, 3, 4]
CASA <115>: antlist[0]
  Out[115]: 0
CASA <116>: antlist[4]
  Out[116]: 4
```

CASA also uses 0-based indexing internally for elements in the Measurement Set (MS – the basic construct that contains visibility and/or single dish data; see Chapter 2). Thus, we will often talk about Field or Antenna “ID”s which will start at 0. For example, the first field in an MS would have `FIELD_ID==0` in the `MSselect` syntax, and can be addressed as be indexed as `field='0'` in most tasks, as well as by name `field='0137+331'` (assuming that’s the name of the first field). You will see these indices in the MS summary from the task `listobs`.

1.3.7.4 Indentation

Python pays attention to the indentation of lines, as it uses indentation to determine the level of nesting in loops. Be careful when cutting and pasting: if you get the wrong indentation, then unpredictable things can happen (usually it just gives an error).

See Appendix B.3 for more information.

1.3.7.5 System shell access

If you want to access system commands from a script, use the `os.system` command (Appendix B.7.1).

In interactive mode, any input line beginning with a `'!'` character is passed verbatim (minus the `'!'`, of course) to the underlying operating system. Also, several common commands (`ls`, `pwd`, `less`) may be executed with or without the `'!'`, although the `cp` command must use `'!'` and `cd` must be executed without the `'!'`. For example:

```
CASA <5>: !rm -r mydata.ms
```

Note that if you want to access a Unix environment variable, you will need to prefix with a double `$$` instead of a single `$` — for example, to print the value of the `$PAGER` variable, you would use

```
CASA <6>: !echo $$PAGER
```

See Appendix B.7 for more information.

1.3.7.6 Executing Python scripts

You can execute Python scripts (ASCII text files containing Python or `casapy` commands) using the `execfile` command. For example, to execute the script contained in the file `myscript.py` (in the current directory), you would type

```
CASA <7>: execfile('myscript.py')
```

or

```
CASA <8>: execfile 'myscript.py'
```

which will invoke the IPython auto-parenthesis feature.

NOTE: in some cases, you can use the IPython `run` command instead, e.g.

```
CASA <9>: run myscript.py
```

In this case, you do not need the quotes around the filename. This is most useful for re-initializing the task parameters, e.g.

```
CASA <10>: run clean.last
```

(see § 1.4.5.8).

See Appendix B.12 for more information.

1.3.8 Getting Help in CASA

1.3.8.1 TAB key

At **any** time, hitting the `<TAB>` key will complete any available commands or variable names and show you a list of the possible completions if there's no unambiguous result. It will also complete filenames in the current directory if no CASA or Python names match.

For example, it can be used to list the available functionality using minimum match; once you have typed enough characters to make the command unique, `<TAB>` will complete it.

```
CASA <15>: cle<TAB>
clean      clean.last  clear      clearcal   clearplot  clearsta
```

1.3.8.2 help <taskname>

Basic information on an application, including the parameters used and their defaults, can be obtained by typing `pdoc task`, `help task`, `help 'task'` or `task?`. The `pdoc task` currently gives the cleanest documentation format with the smallest amount of object-oriented (programmer) output. This inline help provides a one line description of the task and then lists all parameters, a brief description of the parameter, the parameter default, an example setting the parameter and any options if there are limited allowed values for the parameter.

For example:

```
Imports an arbitrary number of VLA archive-format data sets into
a casa measurement set.  If more than one band is present, they
will be put in the same measurement set but in a separate spectral
window.  The task will handle old style and new style VLA (after
July 2007) archive data and apply the tsys to the data and to
the weights.
```



```

Keyword arguments:
archivefiles -- Name of input VLA archive file(s)
    default: none. Must be supplied
    example: archivefiles = 'AP314_A959519.xp1'
    example: archivefiles=['AP314_A950519.xp1','AP314_A950519.xp2']
vis -- Name of output visibility file
    default: none. Must be supplied.
    example: vis='NGC7538.ms'
    Will not over-write existing ms of same name.
    A backup flag-file version 'Original' will be made in
    vis.flagversions. See help flagmanager
bandname -- VLA Frequency band
    default: => '' = all bands
    example: bandname='K'
    Options: '4'=48-96 MHz,'P'=298-345 MHz,'L'=1.15-1.75 GHz,
    'C'=4.2-5.1 GHz,'X'=6.8-9.6 GHz,'U'=13.5-16.3 GHz,
    'K'=20.8-25.8 GHz,'Q'=38-51 GHz
frequencytol -- Tolerance in frequency shift in making spectral windows
    default: => 150000 (Hz). For Doppler shifted data, <10000 Hz may
    may produce too many unnecessary spectral windows.
    example: frequencytol = 1500000.0 (units = Hz)
project -- Project name to import from archive files:
    default: '' => all projects in file
    example: project='AL519'
    project = 'al519' or AL519 will work. Do not include
    leading zeros; project = 'AL0519' will not work.
starttime -- Time after which data will be considered for importing
    default: '' => all: Date must be included.
    syntax: starttime = '2003/1/31/05:05:23'
stoptime -- Time before which data will be considered for importing
    default: '' => all: Date must be included.
    syntax: stoptime = '2003/1/31/08:05:23'
applysys -- Apply data scaling and weight scaling by nominal
    sensitivity (~Tsys)
    default: True. Strongly recommended
autocorr -- import autocorrelations to ms
    default: => False (no autocorrelations)
antnamescheme -- 'old' or 'new' antenna names.
    default => 'new' gives antenna names
    'VA04' or 'EA13 for VLA telescope 04 and 13 (EVLA)
    'old' gives names '04' or '13'
keepblanks -- Should sources with blank names be filled into the data base
    default => false. Do not fill
    These scans are tipping scans (as of June 1, 2009) and should not
    be filled in the visibility data set.

```

You can also get the short help for a CASA tool method by typing 'help tool.method'.

```
CASA <46>: help ia.subimage
```

Summary

Create a (sub)image from a region of the image

Description

This function copies all or part of the image to another on-the-fly Image tool. Both float and complex valued images are supported.

If `{\stfaf outfile}` is given, the subimage is written to the specified disk file. If `{\stfaf outfile}` is unset, the returned Image `ool\` actually references the input image file (i.e. that associated with the Image `ool\` to which you are applying this function). So if you deleted the input image disk file, it would render this `ool\` useless. When you destroy this `ool\` (with the `done` function) the reference connection is broken.

Sometimes it is useful to drop axes of length one (degenerate axes). Use the `{\stfaf dropdeg}` argument if you want to do this.

The output mask is the combination (logical OR) of the default input `\pixelmask\` (if any) and the OTF mask. Any other input `\pixelmask` will not be copied. Use function `need` to copy other masks too.

If the mask has fewer dimensions than the image and if the shape of the dimensions the mask and image have in common are the same, the mask will automatically have the missing dimensions added so it conforms to the image.

If `stretch` is true and if the number of mask dimensions is less than or equal to the number of image dimensions and some axes in the mask are degenerate while the corresponding axes in the image are not, the mask will be stretched in the degenerate dimensions. For example, if the input image has shape `[100, 200, 10]` and the input mask has shape `[100, 200, 1]` and `stretch` is true, the mask will be stretched along the third dimension to shape `[100, 200, 10]`. However if the mask is shape `[100, 200, 2]`, stretching is not possible and an error will result.

Input Parameters:

<code>outfile</code>	Output image file name. Default is unset.
<code>region</code>	Region of interest. Default is whole image.
<code>mask</code>	Mask to use. See help <code>par.mask</code> . Default is none.
<code>dropdeg</code>	Drop degenerate axes false
<code>overwrite</code>	Overwrite (unprompted) pre-existing output file? false
<code>list</code>	List informative messages to the logger true
<code>stretch</code>	Stretch the mask if necessary and possible? false

wantreturn Return an image analysis tool attached to the created subimage true

Example:

```
'''
#
print ' ----      subimage Ex 1      ----'
ia.maketestimage('myfile', overwrite=true)
im2 = ia.subimage()           # a complete copy
r1 = rg.box([10,10],[30,40],[5,5]) # A strided pixel box region
im3 = ia.subimage(outfile='/tmp/foo', region=r1, overwrite=true)
                                     # Explicitly named subimage

im2.done()
im3.done()
ia.close()
#
'''
```

For a full list of keywords associated with the various tools, see the **CASA User Reference Manual**.

1.3.8.3 help and PAGER

Your **PAGER** environment variable (§ 1.3.1) determines how help is displayed in the terminal window where you start CASA. If you set your **bash** environment variable **PAGER=less** (**setenv PAGER less** in **csH**) then typing **help <taskname>** will show you the help but the text will vanish and return you to the command line when you are done viewing it. Setting **PAGER=more** (**setenv PAGER more**) will scroll the help onto your command window and then return you to your prompt (but leaving it on display). Setting **PAGER=cat** (**setenv PAGER cat**) will give you the **more** equivalent without some extra formatting baggage and is the recommended choice.

If you have set **PAGER=more** or **PAGER=less**, the **help** display will be fine, but the display of 'taskname?' will often have confusing formatting content at the beginning (lots of **ESC** surrounding the text). This can be remedied by exiting **casapy** and doing an **'unset PAGER'** (**unsetenv PAGER** in **[t]csH**) at the Unix command line.

You can see the current value of the **PAGER** environment variable with **CASA** by typing:

```
!echo $$PAGER
```

(note the double **\$\$**). This will show what command paging is pointed to.

1.3.8.4 help par.<parameter>

Typing **help par.<parameter>** provides a brief description of a given parameter **<parameter>**.

```
CASA <46>: help par.robust
Help on function robust in module parameter_dictionary:

robust()
  Brigg's robustness parameter.

Options: -2.0 (close to uniform) to 2.0 (close to natural)
```

1.3.8.5 Python help

Typing `help` at the `casapy` prompt with no arguments will bring up the native Python help facility, and give you the `help>` prompt for further information; hitting `<RETURN>` at the help prompt returns you to the `CASA` prompt.

```
CASA <39>: help
-----> help()
```

Welcome to Python 2.7! This is the online help utility.

If this is your first time using Python, you should definitely check out the tutorial on the Internet at <http://docs.python.org/2.7/tutorial/>.

Enter the name of any module, keyword, or topic to get help on writing Python programs and using Python modules. To quit this help utility and return to the interpreter, just type "quit".

To get a list of available modules, keywords, or topics, type "modules", "keywords", or "topics". Each module also comes with a one-line summary of what it does; to list the modules whose summaries contain a given word such as "spam", type "modules spam".

```
help>
```

To exit Python's help, just hit `ENTER`. Further help in working within the Python shell is given in Appendix B.

1.4 Tasks and Tools in CASA

Originally, `CASA` consisted of a collection of tools, combined in the so-called toolkit. Since the majority of prospective users is far more familiar with the concept of tasks, an effort is underway to replace most - if not all - toolkit functionality by tasks.

While running `CASA`, you will have access to and be interacting with tasks, either indirectly by providing parameters to a task, or directly by running a task. Each task has a well defined purpose, and a number of associated parameters, the values of which are to be supplied by the user. Technically speaking, tasks are built on top of tools - when you are running a task, you are running tools in the toolkit, though this should be transparent.

As more tasks are being written, and the functionality of each task is enhanced, there will be less and less reason to run tools in the toolkit. We are working toward a system in which direct access to the underlying toolkit is unnecessary for all standard data processing.

1.4.1 What Tasks are Available?

As mentioned in the introduction, tasks in CASA are python interfaces to the more basic toolkit. Tasks are executed to perform a single job, such as loading, plotting, flagging, calibrating, and imaging the data.

Basic information on tasks, including the parameters used and their defaults, can be obtained by typing `help <taskname>` or `<taskname>?` at the CASA prompt, where `<taskname>` is the name of a given task. As described above in § 1.3.8.2, `help <taskname>` provides a description of the task and then lists all parameters, a brief description of the parameter, the parameter default, an example setting the parameter and any options if there are limited allowed values for the parameter.

To see what tasks are available in CASA, use `tasklist`, e.g.

```
CASA <40>: tasklist
-----> tasklist()
Available tasks, organized by category (experimental tasks in parenthesis ()
  deprecated tasks in curly brackets {}).
```

Import/export	Information	Editing	Manipulation
<code>exportasdm</code>	<code>imhead</code>	<code>fixplanets</code>	<code>concat</code>
<code>exportfits</code>	<code>imreframe</code>	<code>fixvis</code>	<code>conjugatevis</code>
<code>exportuvfits</code>	<code>imstat</code>	<code>flagcmd</code>	<code>cvel</code>
<code>importasdm</code>	<code>imval</code>	<code>flagdata</code>	<code>fixvis</code>
<code>importfits</code>	<code>listcal</code>	<code>flagmanager</code>	<code>hanningsmooth</code>
<code>importfitsidi</code>	<code>listfits</code>	<code>msview</code>	<code>imhead</code>
<code>importmiriad</code>	<code>listhistory</code>	<code>plotms</code>	<code>msmoments</code>
<code>importuvfits</code>	<code>listobs</code>		<code>mstransform</code>
<code>importvla</code>	<code>listpartition</code>		<code>partition</code>
<code>(importevla)</code>	<code>listvis</code>		<code>plotms</code>
<code>(importgmrt)</code>	<code>plotms</code>		<code>split</code>
	<code>plotuv</code>		<code>testconcat</code>
	<code>vishead</code>		<code>uvcontsub</code>
	<code>visstat</code>		<code>virtualconcat</code>
	<code>(asdmsummary)</code>		<code>vishead</code>
	<code>(listsdm)</code>		<code>(cvel2)</code>
	<code>(makemask)</code>		<code>(hanningsmooth2)</code>
			<code>(split2)</code>
			<code>(statwt)</code>
			<code>(uvcontsub3)</code>
Calibration	Modeling	Imaging	Analysis

accum	predictcomp	clean	imcollapse
applycal	setjy	deconvolve	imcontsub
bandpass	uvcontsub	feather	imfit
blcal	uvmodelfit	ft	imhead
calstat	uvsb	imcontsub	immath
clearcal	(uvcontsub3)	(boxit)	immoments
delmod		(csvclean)	impbcor
fixplanets		(tclean)	impv
fluxscale		(widebandpbcor)	imrebin
ft		{mosaic}	imreframe
gaincal		{widefield}	imregrid
gencal			imsmooth
initweights			imstat
listcal			imsubimage
plotants			imtrans
plotbandpass			imval
plotcal			listvis
polcal			rmfit
predictcomp			slsearch
setjy			specsmooth
smoothcal			splattotable
uvmodelfit			(specfit)
uvsb			(spxfit)
wvrgcal			
Visualization	Simulation	Single dish	Utility
-----	-----	-----	-----
clearplot	simanalyze	asap_init	browsetable
imview	simobserve	sdaverage	caltabconvert
msview	(simalma)	sdbaseline	clearplot
plotants		sdbaseline2	clearstat
plotbandpass		sdcal	concat
plotcal		sdcal2	conjugatevis
plotms		sdcoadd	find
plotuv		sdfit	help par.parameter
viewer		sdflag	help taskname
(plotweather)		sdflagmanager	imview
		sdgrid	msview
		sdimaging	plotms
		sdimprocess	rmtables
		sdlist	startup
		sdmath	taskhelp
		sdplot	tasklist
		sdreduce	testconcat
		sdsave	toolhelp
		sdscale	virtualconcat
		sdstat	
		sdtpimaging	
		(tsdbaseline)	
		(tsdcal)	
		(tsdfit)	

(tsdsmoother)

User defined tasks

The tasks with name in parentheses are experimental, those in curly brackets are deprecated and will be removed in future releases. The functionality of deprecated tasks is usually available in some other task (e.g., instead of `mosaic` one should use `clean`). In the above case, the user has not defined any task him/herself.

Typing `taskhelp` provides a one line description of all available tasks.

```
CASA <41>: taskhelp
-----> taskhelp()
Available tasks:
```

```
accum          : Accumulate incremental calibration solutions into a calibration table
applycal       : Apply calibrations solutions(s) to data
asdmsummary    : Summarized description of an ASDM dataset.
autoclean      : CLEAN an image with automatically-chosen clean regions.
bandpass       : Calculates a bandpass calibration solution
blcal          : Calculate a baseline-based calibration solution (gain or bandpass)
boxit          : Box regions in image above given threshold value.
browsetable    : Browse a table (MS, calibration table, image)
calstat        : Displays statistical information on a calibration table
caltabconvert  : Convert old-style caltables into new-style caltables.
clean          : Invert and deconvolve images with selected algorithm
clearcal       : Re-initializes the calibration for a visibility data set
clearplot      : Clear the matplotlib plotter and all layers
clearstat      : Clear all autolock locks
concat         : Concatenate several visibility data sets.
conjugatevis   : Change the sign of the phases in all visibility columns.
csvclean       : This task does an invert of the visibilities and
deconvolve     :
in the image plane.
cvel           : regrid an MS to a new spectral window / channel structure or frame
cvel2          : Regrid an MS or MMS to a new spectral window, channel structure or frame
deconvolve     : Image based deconvolver
delmod         : Deletes model representations in the MS
exportasdm     : Convert a CASA visibility file (MS) into an ALMA or EVLA Science Data Model
exportfits     : Convert a CASA image to a FITS file
exportuvfits   : Convert a CASA visibility data set to a UVFITS file:
feather        : Combine two images using their Fourier transforms
find           : Find string in tasks, task names, parameter names:
fixplanets     : Changes FIELD and SOURCE table entries based on
user-provided
direction or POINTING table, optionally fixes the UVW coordinates
fixvis         : Recalculates (u, v, w) and/or changes Phase Center
flagcmd        : Flagging task based on batches of flag-commands
```

```

flagdata      : All-purpose flagging task based on data-selections
and flagging
modes/algorithms.
flagmanager   : Enable list, save, restore, delete and rename flag version files.
fluxscale    : Bootstrap the flux density scale from standard calibrators
ft           : Insert a source model a visibility set:
gaincal      : Determine temporal gains from calibrator observations
gencal       : Specify Calibration Values of Various Types
hanningsmooth : Hanning smooth frequency channel data to remove Gibbs ringing
hanningsmooth2 : Hanning smooth frequency channel data to remove Gibbs ringing
imcollapse   : Collapse image along one axis, aggregating pixel values along that axis.
imcontsub    : Estimates and subtracts continuum emission from an image cube
imfit        : Fit one or more elliptical Gaussian components on an image region(s)
imhead       : List, get and put image header parameters
immath       : Perform math operations on images
immoments    : Compute moments from an image
impbcor      : Construct a primary beam corrected image from an
image and
a primary beam pattern.
importasdm   : Convert an ALMA Science Data Model observation into a
CASA visibility
file (MS) or single-dish data format (Scantable)
importevla   : Convert an Science Data Model observation into a CASA Measurement Set
importfits   : Convert an image FITS file into a CASA image
importfitsidi : Convert a FITS-IDI file to a CASA visibility data set
importgmrt   : Convert a UVFITS file to a CASA visibility data set
importmiriad : Convert a Miriad visibility file into a CASA MeasurementSet
importuvfits : Convert a UVFITS file to a CASA visibility data set
importvla    : Import VLA archive file(s) to a measurement set
impv        : Construct a position-velocity image by choosing two
points
in the direction plane.
imrebin      : Rebin an image by the specified integer factors
imreframe    : Change the frame in which the image reports its spectral values
imregrid     : regrid an image onto a template image
imsmooth     : Smooth an image or portion of an image
imstat       : Displays statistical information from an image or image region
imsubimage   : Create a (sub)image from a region of the image
imtrans      : Reorder image axes
imval        : Get the data value(s) and/or mask value in an image.
imview       : View an image
initweights  : Initializes weight information in the MS
listcal      : List antenna gain solutions
listfits     : List the HDU and typical data rows of a fits file:
listhistory  : List the processing history of a dataset:
listobs      : List the summary of a data set in the logger or in a file
listpartition : List the summary of a multi-MS data set in the logger or in a file
listsdm      : Lists observation information present in an SDM directory.
listvis      : List measurement set visibilities.
makemask     : Makes and manipulates image masks
mosaic       : Create a multi-field deconvolved image with selected algorithm

```



```

msmoments      : Compute moments from an MS
mstransform    : Split the MS, combine/separate/regrid spws and do
channel
  and time averaging
msuvbin        : grid the visibility data onto a defined uniform grid
(in the form of an ms);
multiple MS's can be done onto the same grid
msview         : View a visibility data set
partition      : Task to produce Multi-MSs using parallelism
pclean         : Invert and deconvolve images with parallel engines
plotants       : Plot the antenna distribution in the local reference frame:
plotbandpass   : Makes detailed plots of Tsys and bandpass solutions.
plotcal        : An all-purpose plotter for calibration results
plotms         : A plotter/interactive flagger for visibility data.
plotuv         : Plot the baseline distribution
plotweather    : Plot elements of the weather table; estimate opacity.
polcal         : Determine instrumental polarization calibrations
predictcomp    : Make a component list for a known calibrator
rmfit          : Calculate rotation measure.
rmtables       : Remove tables cleanly, use this instead of rm -rf
sdaverage      : ASAP SD task: averaging and smoothing of spectra
sdbaseline     : Fit/subtract a spectral baseline
sdbaseline2    : Fit/subtract a spectral baseline
sdcalf         : ASAP SD calibration task
sdcalf2        : ASAP SD calibration task
sdcoadd        : Coadd multiple scantables into one
sdfit          : Fit a spectral line
sdflag         : ASAP SD spectral spectral/row flagging task
sdflagmanager  : ASAP SD task to manipulate flag version files
sdgrid         : SD gridding task
sdimaging      : SD task: imaging for total power and spectral data
sdimprocess    : Task for single-dish image processing
sdlist         : list summary of single dish data
sdmath         : ASAP SD task for simple arithmetic of spectra
sdplot         : ASAP SD plotting task
sdreduce       : ASAP SD task: do sdcalf, sdaverage, and sdbaseline in one task
sdsave         : Save the sd spectra in various format
sdscale        : Scale the sd spectra
sdstat         : list statistics of spectral
sdtpimaging    : SD task: do a simple calibration (baseline
subtraction) and
imaging for total power data
setjy          : Fills the model column with the visibilities of a calibrator
simalma        : Simulation task for ALMA
simanalyze     : image and analyze measurement sets created with simobserve
simobserve      : visibility simulation task
slsearch       : Search a spectral line table.
smoothcal      : Smooth calibration solution(s) derived from one or more sources:
specfit        : Fit 1-dimensional gaussians and/or polynomial models
to an image or
image region

```

```

specsmooth      : Smooth an image region in one dimension
splattotable    : Convert a downloaded Splatalogue spectral line list to a casa table.
split           : Create a visibility subset from an existing visibility set
split2          : Create a visibility subset from an existing visibility set
spxfit          : Fit a 1-dimensional model(s) to an image(s) or region
for
determination of spectral index.
ssoflux         : Fills the model column with the visibilities of a calibrator
statwt          : Reweight visibilities according to their scatter (Experimental)
tclean          : Radio Interferometric Image Reconstruction
testconcat      : Concatenate the subtables of several visibility data
sets,
not the MAIN bulk data.
tsdbaseline     : Fit/subtract a spectral baseline
tsdcal          : MS SD calibration task
tsdfit          : Fit a spectral line
tsdsmooth       : Smooth spectral data
uvcontsub       : Continuum fitting and subtraction in the uv plane
uvcontsub3      : An experimental clone of uvcontsub
uvmodelfit      : Fit a single component source model to the uv data
uvsb            : Subtract/add model from/to the corrected visibility data.
viewer          : View an image or visibility data set
virtualconcat   : Concatenate several visibility data sets into a multi-MS
vishead         : List, summary, get, and put metadata in a measurement set
visstat         : Displays statistical information from a Measurement
Set,
or from a Multi-MS
widebandpbcor  : Wideband PB-correction on the output of the MS-MFS algorithm
widefield       : Wide-field imaging and deconvolution with selected algorithm
wvrgcal         : Generate a gain table based on Water Vapour Radiometer data

```

Typing `startup` will provide the startup page displayed when entering CASA. The startup screen lists the various options to obtain help within CASA.

```

CASA <26>: startup
-----> startup()

```

```

-----
For help use the following commands:
tasklist          - Task list organized by category
taskhelp          - One line summary of available tasks
help taskname     - Full help for task
toolhelp          - One line summary of available tools
help par.parametername - Full help for parameter name
-----

```

1.4.2 Running Tasks and Tools

Tools are functions linked to the Python interface which must be called by name with arguments. Tasks have higher-level capabilities than tools. Tasks require input parameters which maybe be specified when you call the task as a function, or be set as parameters in the interface. A task, like a tool, is a function under Python and may be written in Python, C, or C++ (the CASA toolkit is made up of C++ functions).

There are two distinct ways to run tasks. You can either set the global CASA parameters relevant to the task and tell the task to “go”, or you can call the task as a function with one or more arguments specified. These two invocation methods differ in whether the global parameter values are used or not.

For example,

```
default('plotxy')
vis='ngc5921.ms'
xaxis='channel'
yaxis='amp'
datacolumn='data'
go
```

will execute `plotxy` with the set values for the parameters (see § 1.4.5). Instead of using `go` command (§ 1.4.5.3) to invoke the task, you can also call the task with no arguments, e.g.

```
default('plotxy')
vis='ngc5921.ms'
xaxis='channel'
yaxis='amp'
datacolumn='data'
plotxy()
```

which will also use the global parameter values.

Second, one may call tasks and tools by name with parameters set on the same line. Parameters may be set either as explicit `<parameter>=<value>` arguments, or as a series of comma delimited `<value>`s in the correct order for that task or tool. Note that missing parameters will *use the default values for that task*. For example, the following are equivalent:

```
# Specify parameter names for each keyword input:
plotxy(vis='ngc5921.ms',xaxis='channel',yaxis='amp',datacolumn='data')
# when specifying the parameter name, order doesn't matter, e.g.:
plotxy(xaxis='channel',vis='ngc5921.ms',datacolumn='data',yaxis='amp')
# use parameter order for invoking tasks
plotxy('ngc5921.ms','channel','amp','data')
```

This non-use of globals when calling as a function is so that robust scripts can be written. One need only cut-and-paste the calls and need not worry about the state of the global variables or

what has been run previously. It is also more like the standard behavior of function calls in Python and other languages.

Tools can only be called in this second manner by name, with arguments (§ 1.4.6). Tools never use the global parameters and the related mechanisms of `inp` and `go`.

1.4.2.1 Aborting Synchronous Tasks

If you are running CASA tasks synchronously, then you can usually use `CTRL-C` to abort execution of the task. If this does not work, try `CTRL-Z` followed by a `kill`. See § 1.3.5 for more on these methods to abort CASA execution.

You may have to quit and restart CASA after an abort, as the internal state can get mixed up.

1.4.3 Getting Return Values

Some tasks and tools return a record (usually a Python dictionary) to the interface. For example, the `imstat` task (§ 6.9) returns a dictionary with the image statistics in it. To catch these return values into a Python variable, you **MUST** assign that variable to the task call, e.g.

```
xstat = imstat('ngc5921.clean.image')
```

or

```
default('imstat')
imagenname = 'ngc5921.clean.image'
xstat = imstat()
```

Note that tools that return values work in the same way (§ 1.4.6).

You can print or use the return value in Python for controlling scripts. For example,

```
CASA <1>: xstat = imstat('ngc5921.clean.image')
CASA <2>: xstat
Out [2]:
{'blc': array([0, 0, 0, 0]),
 'blcf': '15:24:08.404, +04.31.59.181, I, 1.41281e+09Hz',
 'flux': array([ 4.15292207]),
 'max': array([ 0.05240594]),
 'maxpos': array([134, 134,  0,  38]),
 'maxposf': '15:21:53.976, +05.05.29.998, I, 1.41374e+09Hz',
 'mean': array([ 1.62978083e-05]),
 'medabsdevmed': array([ 0.00127287]),
 'median': array([ -1.10467618e-05]),
 'min': array([-0.0105249]),
 'minpos': array([160,  1,  0,  30]),
 'minposf': '15:21:27.899, +04.32.14.923, I, 1.41354e+09Hz',
```

```
'npts': array([ 3014656.]),
'quartile': array([ 0.00254587]),
'rms': array([ 0.00201818]),
'sigma': array([ 0.00201811]),
'sum': array([ 49.1322855]),
'sumsq': array([ 12.27880404]),
'trc': array([255, 255, 0, 45]),
'trcf': '15:19:52.390, +05.35.44.246, I, 1.41391e+09Hz'}
CASA <3>: myrms = xstat['rms'][0]
CASA <4>: print 10.0*myrms
0.0201817648485
```

If you do not catch the return variable, it will be lost

```
imstat('ngc5921.clean.image')
```

or

```
default('imstat')
imagenname = 'ngc5921.clean.image'
imstat()
```

and spewed to terminal. Note that `go` will trap and lose the return value, e.g.

```
default('imstat')
imagenname = 'ngc5921.clean.image'
go
```

will not dump the return to the terminal either.

1.4.4 Running Tasks Asynchronously

By default, most tasks run synchronously in the foreground. Many tasks, particularly those that can take a long time to execute, have the `async` parameter. This allows the user to send the task to the background for execution.

1.4.5 Setting Parameters and Invoking Tasks

One can set parameters for tasks (but not for tools) by performing the assignment within the CASA shell and then inspecting them using the `inp` command:

```
CASA <30>: default(bandpass)
CASA <31>: vis = 'ngc5921.demo.ms'
CASA <32>: caltable = 'ngc5921.demo.bcal'
CASA <33>: field = '0'
```

Inside the Toolkit:

In the current version of CASA, you cannot use the task parameter setting features, such as the `inp`, `default`, or `go` commands, for the tools.

```

CASA <34>: refant = '15'
CASA <35>: inp('bandpass')
# bandpass :: Calculates a bandpass calibration solution
vis                = 'ngc5921.demo.ms' # Name of input visibility file
caltable           = 'ngc5921.demo.bcal' # Name of output gain calibration
# table
field              = '0'                # Select field using field id(s) or
# field name(s)
spw                = ''                 # Select spectral window/channels
intent             = ''                 # Select observing intent
selectdata         = True               # Other data selection parameters
  timerange        = ''                 # Select data based on time range
  uvrange          = ''                 # Select data within uvrange (default
# units meters)
  antenna          = ''                 # Select data based on antenna/baseline
  scan             = ''                 # Scan number range
  observation       = ''                 # Select by observation ID(s)
  msselect         = ''                 # Optional complex data selection
# (ignore for now)

solint             = 'inf'               # Solution interval in time[,freq]
combine            = 'scan'              # Data axes which to combine for solve
# (obs, scan, spw, and/or field)
refant             = '15'                # Reference antenna name(s)
minblperant        = 4                  # Minimum baselines _per antenna_
# required for solve
minsnr             = 3.0                # Reject solutions below this SNR (only
# applies for bandtype = B)
solnorm            = False               # Normalize average solution amplitudes
# to 1.0
bandtype           = 'B'                 # Type of bandpass solution (B or
# BPOLY)
  fillgaps         = 0                  # Fill flagged solution channels by
# interpolation

smodel             = []                 # Point source Stokes parameters for
# source model.
append             = False               # Append solutions to the (existing)
# table
docallib           = False               # Use callib or traditional cal apply
# parameters
  gaintable        = []                 # Gain calibration table(s) to apply on
# the fly
  gainfield        = []                 # Select a subset of calibrators from
# gaintable(s)
  interp           = []                 # Interpolation mode (in time) to use
# for each gaintable
  spwmap           = []                 # Spectral windows combinations to form
# for gaintables(s)

parang             = False               # Apply parallactic angle correction

```

See § 1.4.5.4 below for more details on the use of the `inputs` command.

All task parameters have **global** scope within CASA: the parameter values are common to all tasks and also at the CASA command line. This allows the convenience of not changing parameters that are shared between tasks but does require care when chaining together sequences of task invocations (to ensure proper values are provided).

If you want to reset the input keywords for a single task, use the `default` command (§ 1.4.5.2). For example, to set the defaults for the `bandpass` task, type:

```
CASA <30>: default('bandpass')
```

as we did above.

To inspect a single parameter value just type it at the command line. Continuing the above example:

```
CASA <36>: combine
  Out[14]: 'scan'
```

CASA parameters are just Python variables.

Parameters for a given task can be saved by using the `saveinputs` command (see § 1.4.5.5) and restored using the `execfile '<filename>'` command. Note that if the task is successfully executed, then a `<taskname>.last` file is created in the working directory containing the parameter values (see § 1.4.5.8).

We now describe the individual CASA task parameter interface commands and features in more detail.

1.4.5.1 The scope of parameters in CASA

All task parameters have **global** scope within CASA: the parameter values are common to all tasks and also at the CASA command line. This allows the convenience of not changing parameters that are shared between tasks but does require care when chaining together sequences of task invocations (to ensure proper values are provided). Tasks **DO NOT** change the values of the global parameters, nor does the invocation of tasks using the functional call with arguments change the globals.

This does mean that unless you do an explicit `default` of the task (§ 1.4.5.2), previously set values may be unexpectedly used if you do not inspect the `inp` carefully. For example, good practice is:

Advanced Tip

By default, the scope of CASA parameters is global, as stated here. However, if you call a task as a function with one or more arguments specified, e.g. `task(arg1=val1, ...)`, then non-specified parameters will be defaulted and no globals used. This makes scripting more robust. Tasks **DO NOT** change the value of globals.

```

default('imhead')
imagenname = 'ngc5921.demo.cleanimg.image'
mode = 'list'
imhead()

```

If you supply the task call with arguments, then these will be used for the values of those parameters (see above). However, if some but not all arguments are supplied, then those parameters not given as arguments will default and NOT use the current global values. Thus,

```
imhead('ngc5921.demo.cleanimg.image',mode='list')
```

will reproduce the above.

1.4.5.2 The default Command

Each task has a special set of default parameters defined for its parameters. You can use the `default` command to reset the parameters for a specified task (or the current task as defined by the `taskname` variable) to their default.

Important Note: The `default` command resets the values of the task parameters to a set of “defaults” as specified in the task code. Some defaults are blank strings `''` or empty lists `[]`, others are specific numerical values, strings, or lists. It is important to understand that just setting a string parameter to an empty string `''` is not setting it to its default! Some parameters do not have a blank as an allowed value. See the `help` for a particular task to find out its default. If `''` is the default or an allowed value, it will say so explicitly.

For example, suppose we have been running CASA on a particular dataset, e.g.

```

CASA <40>: inp clean
-----> inp('clean')
# clean :: Deconvolve an image with selected algorithm
vis                = 'ngc5921.demo.src.split.ms.contsub' # name of input visibility file
imagenname         = 'ngc5921.demo.cleanimg' # Pre-name of output images
field              = '0' # Field Name
spw                = '' # Spectral windows:channels: '' is all
selectdata         = False # Other data selection parameters
mode               = 'channel' # Type of selection (mfs, channel, velocity, frequency)
    nchan          = 46 # Number of channels (planes) in output image
    start          = 5 # first input channel to use
    width          = 1 # Number of input channels to average
    interpolation   = 'nearest' # Spectral interpolation (nearest, linear, cubic)
niter              = 6000 # Maximum number of iterations
...

```

and now we wish to switch to a different one. We can reset the parameter values using `default`:


```

CASA <41>: default
-----> default()

CASA <42>: inp
-----> inp()
# clean :: Deconvolve an image with selected algorithm
vis          =      ''          # name of input visibility file
imagename    =      ''          # Pre-name of output images
field        =      ''          # Field Name
spw          =      ''          # Spectral windows:channels: '' is all
selectdata   =      False       # Other data selection parameters
mode         =      'mfs'       # Type of selection (mfs, channel, velocity, frequency)
niter        =      500         # Maximum number of iterations
...

```

It is good practice to use `default` before running a task if you are unsure what state the CASA global variables are in.

ALERT: You currently can only reset ALL of the parameters for a given task to their defaults. In an upcoming update we will allow the `default` command to take a second argument with a specific parameter to default its value.

1.4.5.3 The go Command

You can execute a task using the `go` command, either explicitly

```

CASA <44>: go listobs
-----> go(listobs)
Executing: listobs()
...

```

or implicitly if `taskname` is defined (e.g. by previous use of `default` or `inp`)

```

CASA <45>: taskname = 'clean'
CASA <46>: go
-----> go()
Executing: clean()
...

```

You can also execute a task simply by typing the `taskname`.

```

CASA <46>: clean
-----> clean()
Executing: clean()
...

```

The `go` command can also be used to launch a different task without changing the current `taskname`, without disrupting the `inp` process on the current task you are working on. For example

```

default 'gaincal' # set current task to gaincal and default
vis = 'n5921.ms' # set the working ms
...           # set some more parameters
go listobs    # launch listobs w/o changing current task
inp          # see the inputs for gaincal (not listobs!)

```

ALERT: Doing `go listobs(vis='foo.ms')` will currently change the `taskname`, and will change `vis`, which might not be what is desired.

1.4.5.4 The `inp` Command

You can set the values for the parameters for tasks (but currently not for tools) by performing the assignment within the CASA shell and then inspecting them using the `inp` command. This command can be invoked in any of three ways: via function call `inp('<taskname>')` or `inp(<taskname>)`, without parentheses `inp '<taskname>'` or `inp <taskname>`, or using the current `taskname` variable setting with `inp`. For example,

```

CASA <1>: inp('clean')
...
CASA <2>: inp 'clean'
-----> inp('clean')
...
CASA <3>: inp(clean)
...
CASA <4>: inp clean
-----> inp(clean)
...
CASA <5>: taskname = 'clean'
CASA <6>: inp
-----> inp()

```

all do the same thing.

When you invoke the task inputs via `inp`, you see a list of the parameters, their current values, and a short description of what that parameters does. For example, starting from the default values,

```

CASA <18>: inp('clean')
# clean :: Deconvolve an image with selected algorithm
vis          =          ''          # name of input visibility file
imagename    =          ''          # Pre-name of output images
field        =          ''          # Field Name
spw          =          ''          # Spectral windows:channels: '' is all
selectdata   =      False          # Other data selection parameters
mode         =      'mfs'          # Type of selection (mfs, channel, velocity, frequency)
niter        =          500         # Maximum number of iterations
gain         =          0.1         # Loop gain for cleaning
threshold    =      '0.0mJy'       # Flux level to stop cleaning. Must include units
psfmode      =      'clark'        # method of PSF calculation to use during minor cycles
imagermode   =          ''          # Use csclean or mosaic. If '', use psfmode

```

```

multiscale      =      []      # multi-scale deconvolution scales (pixels)
interactive     =      False   # use interactive clean (with GUI viewer)
mask           =      []      # cleanbox(es), mask image(s), and/or region(s)
imsize        = [256, 256]    # x and y image size in pixels
cell          = ['1.0arcsec', '1.0arcsec'] # x and y cell size. default unit arcsec
phasecenter    =      ''      # Image phase center: position or field index
restfreq      =      ''      # rest frequency to assign to image (see help)
stokes        =      'I'     # Stokes params to image (eg I,IV, QU,IQUV)
weighting     = 'natural'    # Weighting of uv (natural, uniform, briggs, ...)
uvtaper       =      False   # Apply additional uv tapering of visibilities.
modelimage    =      ''      # Name of model image(s) to initialize cleaning
restoringbeam =      ['']    # Output Gaussian restoring beam for CLEAN image
pbcor        =      False   # Output primary beam-corrected image
minpb        =      0.1     # Minimum PB level to use

```

Figure 1.1 shows how this will look to you on your terminal. Note that some parameters are in boldface with a gray background. This means that some values for this parameter will cause it to *expand*, revealing new *sub-parameters* to be set.

```

CASA <2>: default('clean')
CASA <3>: inp('clean')
# clean :: Deconvolve an image with selected algorithm
vis          =      ''      # name of input visibility file
imagename    =      ''      # Pre-name of output images
field        =      ''      # Field Name
spw          =      ''      # Spectral windows:channels: '' is all
selectdata  =      False   # Other data selection parameters
mode       =      'mfs'    # Type of selection (mfs, channel, velocity, frequency)
niter        =      300     # Maximum number of iterations
gain         =      0.1     # Loop gain for cleaning
threshold    = '0.0mJy'    # Flux level to stop cleaning. Must include units
psfmode     =      'clark'  # method of PSF calculation to use during minor cycles
imagermode =      ''      # Use csclean or mosaic. If '', use psfmode
multiscale =      []      # set deconvolution scales (pixels), default: multiscale=[] (standard CLEAN)
interactive =      False   # use interactive clean (with GUI viewer)
mask         =      []      # cleanbox(es), mask image(s), and/or region(s) used in cleaning
imsize      = [256, 256]    # x and y image size in pixels, symmetric for single value
cell        = ['1.0arcsec', '1.0arcsec'] # x and y cell size. default unit arcsec
phasecenter =      ''      # Image phase center: position or field index
restfreq    =      ''      # rest frequency to assign to image (see help)
stokes      =      'I'     # Stokes params to image (eg I,IV, QU,IQUV)
weighting  = 'natural'    # Weighting of uv (natural, uniform, briggs, ...)
uvtaper   =      False   # Apply additional uv tapering of visibilities.
modelimage  =      ''      # Name of model image(s) to initialize cleaning
restoringbeam =      ['']  # Output Gaussian restoring beam for CLEAN image
pbcor      =      False   # Output primary beam-corrected image
minpb      =      0.1     # Minimum PB level to use
async      =      False   # If true the taskname must be started using clean(...)
CASA <4>:

```

Figure 1.1: Screen shot of the default CASA inputs for task `clean`.

CASA uses color and font to indicate different properties of parameters and their values:

Parameter and Values in CASA inp

	Text Font	Text Color	Highlight	Indentation	Meaning
Parameters:					
	plain	black	none	none	standard parameter
	bold	black	grey	none	expandable parameter
	plain	green	none	yes	sub-parameter
Values:					
	plain	black	none	none	default value
	plain	blue	none	none	non-default value
	plain	red	none	none	invalid value

Figure 1.2 shows what happens when you set some of the `clean` parameters to non-default values. Some have opened up sub-parameters, which can now be seen and set. Figure 1.3 shows what happens when you set a parameter, in this case `vis` and `mode`, to an invalid value. Its value now appears in red. Reasons for invalidation include incorrect type, an invalid menu choice, or a filename that does not exist. For example, since `vis` expects a filename, it will be invalidated (red) if it is set to a non-string value, or a string that is not the name of a file that can be found. The `mode='happy'` is invalid because it's not a supported choice ('mfs', 'channel', 'velocity', or 'frequency').

```

xterm <5>
CASA <4>: tget('clean')
Restored parameters from file clean.last

CASA <5>: inp('clean')
# clean :: Deconvolve an image with selected algorithm
vis          = 'ngc5921.demo.src.split.ms.contsub' # name of input visibility file
image_name   = 'ngc5921.demo.cleaning' # Pre-name of output images
field        = '' # Field Name
spw          = '' # Spectral windows:channels: '' is all
selectdata   = False # Other data selection parameters
mode        = 'channel' # Type of selection (mfs, channel, velocity, frequency)
  nchan      = 46 # Number of channels (planes) in output image
  start      = 5 # first input channel to use
  width      = 1 # Number of input channels to average
  interpolation = 'nearest' # Type of spectral interpolation of visibilities (nearest, linear, cubic)

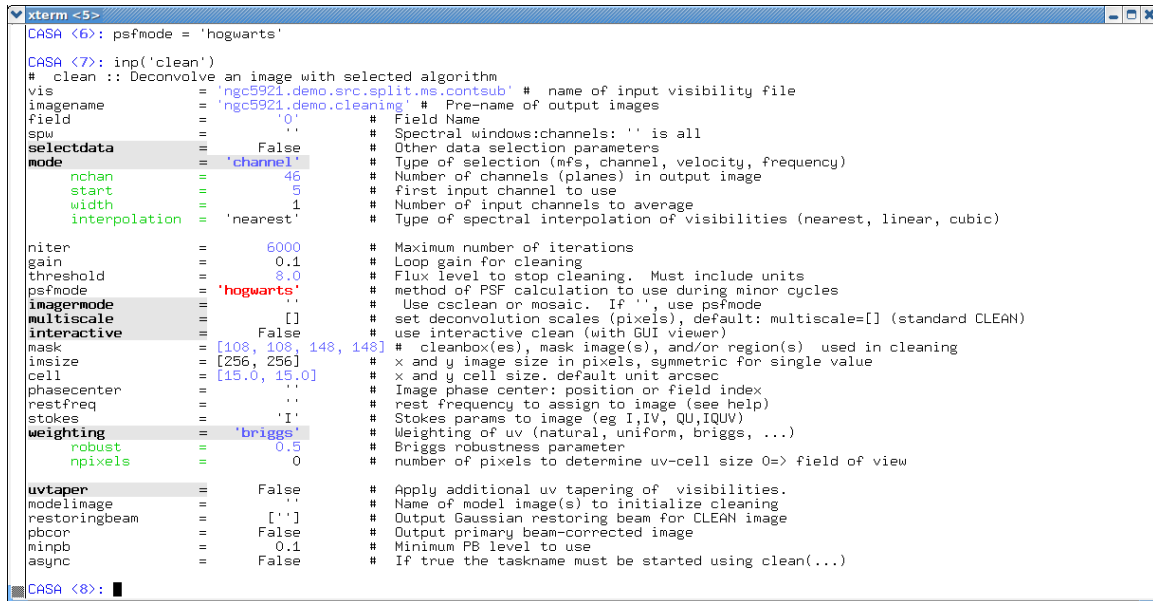
niter        = 6000 # Maximum number of iterations
gain         = 0.1 # Loop gain for cleaning
threshold    = 8.0 # Flux level to stop cleaning. Must include units
psfmode      = 'hogbom' # method of PSF calculation to use during minor cycles
imagermode  = '' # Use csclean or mosaic. If '', use psfmode
multiscale = [] # set deconvolution scales (pixels), default: multiscale=[] (standard CLEAN)
interactive = False # use interactive clean (with GUI viewer)
mask         = [108, 108, 148, 148] # cleanbox(es), mask image(s), and/or region(s) used in cleaning
imsz         = [256, 256] # x and y image size in pixels, symmetric for single value
cell         = [15.0, 15.0] # x and y cell size. default unit arcsec
phasecenter  = '' # Image phase center: position or field index
restfreq     = '' # rest frequency to assign to image (see help)
stokes       = 'I' # Stokes params to image (eg I,IV,QU,IQUV)
weighting   = 'briggs' # Weighting of uv (natural, uniform, briggs, ...)
  robust     = 0.5 # Briggs robustness parameter
  npixels    = 0 # number of pixels to determine uv-cell size 0=> field of view

uvtaper    = False # Apply additional uv tapering of visibilities.
modelimage   = '' # Name of model image(s) to initialize cleaning.
restoringbeam = [] # Output Gaussian restoring beam for CLEAN image
pbcor        = False # Output primary beam-corrected image
minpb        = 0.1 # Minimum PB level to use
async        = False # If true the taskname must be started using clean(...)

CASA <6>:

```

Figure 1.2: The `clean` inputs after setting values away from their defaults (blue text). Note that some of the boldface ones have opened up new dependent sub-parameters (indented and green).



```

xterm <5>
CASA <6>: psfmode = 'hogwarts'
CASA <7>: inp('clean')
# clean :: Deconvolve an image with selected algorithm
vis          = 'ngc5921.demo.src.split.ms.contsub' # name of input visibility file
imagename    = 'ngc5921.demo.cleaning' # Pre-name of output images
field        = '0' # Field Name
spw          = '' # Spectral windows:channels: '' is all
selectdata   = False # Other data selection parameters
mode         = 'channel' # Type of selection (mfs, channel, velocity, frequency)
  nchan      = 46 # Number of channels (planes) in output image
  start      = 5 # first input channel to use
  width      = 1 # Number of input channels to average
  interpolation = 'nearest' # Type of spectral interpolation of visibilities (nearest, linear, cubic)

niter        = 6000 # Maximum number of iterations
gain         = 0.1 # Loop gain for cleaning
threshold    = 8.0 # Flux level to stop cleaning. Must include units
psfmode      = 'hogwarts' # method of PSF calculation to use during minor cycles
imagermode   = '' # Use csclean or mosaic. If '', use psfmode
multiscale   = [] # set deconvolution scales (pixels), default: multiscale=[] (standard CLEAN)
interactive  = False # use interactive clean (with GUI viewer)
mask         = [108, 108, 148, 148] # cleanbox(es), mask image(s), and/or region(s) used in cleaning
imsz         = [256, 256] # x and y image size in pixels, symmetric for single value
cell         = [15.0, 15.0] # x and y cell size. default unit arcsec
phasecenter  = '' # Image phase center: position or field index
restfreq     = '' # rest frequency to assign to image (see help)
stokes       = 'I' # Stokes params to image (eg I,IV,QU,IQUV)
weighting    = 'briggs' # Weighting of uv (natural, uniform, briggs, ...)
  robust     = 0.5 # Briggs robustness parameter
  npixels    = 0 # number of pixels to determine uv-cell size 0=> field of view

uvtaper      = False # Apply additional uv tapering of visibilities.
modelimage   = '' # Name of model image(s) to initialize cleaning.
restoringbeam = [] # Output Gaussian restoring beam for CLEAN image
pbcor        = False # Output primary beam-corrected image
minpb        = 0.1 # Minimum PB level to use
async        = False # If true the taskname must be started using clean(...)

CASA <8>:

```

Figure 1.3: The `clean` inputs where one parameter has been set to an invalid value. This is drawn in red to draw attention to the problem. This hapless user probably confused the `'hogbom'` clean algorithm with Harry Potter.

1.4.5.5 The `saveinputs` Command

The `saveinputs` command will save the current values of a given task parameters to a Python (plain ascii) file. It can take up to two arguments, e.g.

```
saveinputs(taskname, outfile)
```

The first is the usual `taskname` parameter. The second is the name for the output Python file. If there is no second argument, for example,

```
saveinputs('clean')
```

a file with name `<taskname>.saved` (in this case `'clean.saved'` will be created or overwritten if extant. If invoked with no arguments, e.g.

```
saveinputs
```

it will use the current values of the `taskname` variable (as set using `inp <taskname>` or default `<taskname>`). You can also use the `taskname` global parameter explicitly,

```
saveinputs(taskname, taskname+'_1.save')
```

For example, starting from default values

```
CASA <1>: default('listobs')
CASA <2>: vis='ngc5921.demo.ms'
CASA <3>: saveinputs
CASA <4>: !more 'listobs.saved'
taskname      = "listobs"
vis           = "ngc5921.demo.ms"
selectdata    = True
spw           = ""
field         = ""
antenna       = ""
uvrange       = ""
timerange     = ""
correlation   = ""
scan          = ""
intent        = ""
feed          = ""
array         = ""
observation   = ""
verbose       = True
listfile      = ""
#listobs(vis="ngc5921.demo.ms",selectdata=True,spw="",field="",
antenna="",uvrange="",timerange="",correlation="",scan="",intent="",
feed="",array="",observation="",verbose=True,listfile="")
```

To read these back in, use the Python `execfile` command. For example,

```
CASA <5>: execfile('listobs.saved')
```

and we are back.

An example save to a custom named file:

```
CASA <6>: saveinputs('listobs','ngc5921_listobs.par')
```

You can also use the CASA `tget` command (see § 1.4.5.6 below) instead of the Python `execfile` to restore your inputs.

1.4.5.6 The `tget` Command

The `tget` command will recover saved values of the inputs of tasks. This is a convenient alternative to using the Python `execfile` command (see above).

Typing `tget` without a taskname will recover the saved values of the inputs for the current task as given in the current value of the `taskname` parameter.

Adding a task name, e.g. `tget <taskname>` will recover values for the specified task. This is done by searching for 1) a `<taskname>.last` file (see § 1.4.5.8 below), then for 2) a `<taskname>.saved` file (see § 1.4.5.5 above), and then executing the Python in these files.

For example,

```
default('gaincal') # set current task to gaincal and default
tget               # read saved inputs from gaincal.last (or gaincal.saved)
inp               # see these inputs!
tget bandpass     # now get from bandpass.last (or bandpass.saved)
inp               # task is now bandpass, with recovered inputs
```

1.4.5.7 The tput Command

The `tput` command will save the current parameter values of a task to its `<taskname>.last` file. This is a shorthand to `saveinputs` and is a counterpart to `tget`.

Typing `tput` without a taskname will save the values of the inputs for the current task as given in the current value of the `taskname` parameter.

Adding a task name, e.g. `tget <taskname>` will save the values for the specified task.

For example,

```
default('gaincal') # set current task to gaincal and default
tget               # read saved inputs from gaincal.last (or gaincal.saved)
inp               # see these inputs!
vis = 'new.ms'    # change the vis parameter
tput               # save back to the gaincal.last file for later use
```

1.4.5.8 The .last file

Whenever you successfully execute a CASA task, a Python script file called `<taskname>.last` will be written (or over-written) into the current working directory. For example, if you ran the `listobs` task as detailed above, then

```
CASA <14>: vis = 'ngc5921.ms'

CASA <15>: verbose = True

CASA <16>: listobs()

CASA <17>: !more 'listobs.last'
IPython system call: more listobs.last
taskname      = "listobs"
vis           = "ngc5921.ms"
verbose       = True
listfile      = ""
#listobs(vis="ngc5921.ms",verbose=False,listfile="")
```

You can restore the parameter values from the save file using

```
CASA <18>: execfile('listobs.last')
```

or

```
CASA <19>: run listobs.last
```

Note that the `.last` file is generally not created until the task actually finished (successfully), so it is often best to manually create a save file beforehand using the `saveinputs` command if you are running a critical task that you strongly desire to have the inputs saved for.

1.4.6 Tools in CASA

The CASA *toolkit* is the foundation of the functionality in the package, and consists of a suite of functions that are callable from Python. The tools are used by the tasks, and can be used by advanced users to perform operations that are not available through the tasks.

It is beyond the scope of this reference to describe the toolkit in detail. Occasionally, examples will be given that utilize the tools (e.g. § 6.25). In short, tools are always called as functions, with any parameters that are not to be defaulted given as arguments. For example:

```
ia.open('ngc5921.chan21.clean.cleanbox.mask')
ia.calcmask('"ngc5921.chan21.clean.cleanbox.mask">0.5', 'mymask')
ia.summary()
ia.close()
```

uses the `image` tool (`ia`) to turn a clean mask image into an image mask. Tools never use the CASA global parameters.

To find what tools are available, use the `toolhelp` command:

```
CASA <4>: toolhelp
-----> toolhelp()
```

Available tools:

```
af : Agent flagger utilities
at : Juan Pardo ATM library
ca : Calibration analysis utilities
cb : Calibration utilities
cl : Component list utilities
cp : Cal solution plotting utilities
cs : Coordinate system utilities
cu : Class utilities
dc : Deconvolver utilities
fi : Fitting utilities
fn : Functional utilities
```



```

ia : Image analysis utilities
im : Imaging utilities
lm : linear mosaic
me : Measures utilities
ms : MeasurementSet (MS) utilities
msmd : MS metadata accessors
mt : MS transformer utilities
qa : Quanta utilities
pm : PlotMS utilities
po : Imagepol utilities
rg : Region manipulation utilities
sl : Spectral line import and search
sm : Simulation utilities
tb : Table utilities (selection, extraction, etc)
tp : Table plotting utilities
vp : Voltage pattern/primary beam utilities
---
pl : pylab functions (e.g., pl.title, etc)
sd : Single dish utilities
---
```

You can find much more information about the toolkit in the *CASA User Reference Manual*:

<http://casa.nrao.edu/docs/casaref/CasaRef.html>

1.5 Getting the most out of CASA

There are some other general things you should know about using CASA in order to make things go smoothly during your data reduction.

1.5.1 Your command line history

Your command line history is automatically maintained and stored as `ipython.log` in your local directory. This file can be edited and re-executed as appropriate using the `execfile '<filename>'` feature.

You can also use the “up-arrow” and “down-arrow” keys for command line recall in the `casapy` interface. If you start typing text, and then use “up-arrow”, you will navigate back through commands matching what you typed.

1.5.2 Logging your session

The output from CASA commands is sent to the file `casapy-YYYYMMDD-HHMMSS.log` in your local directory, where YYYYMMDD-HHMMSS are the UT date and time when CASA was started up. New starts of CASA create new log files.

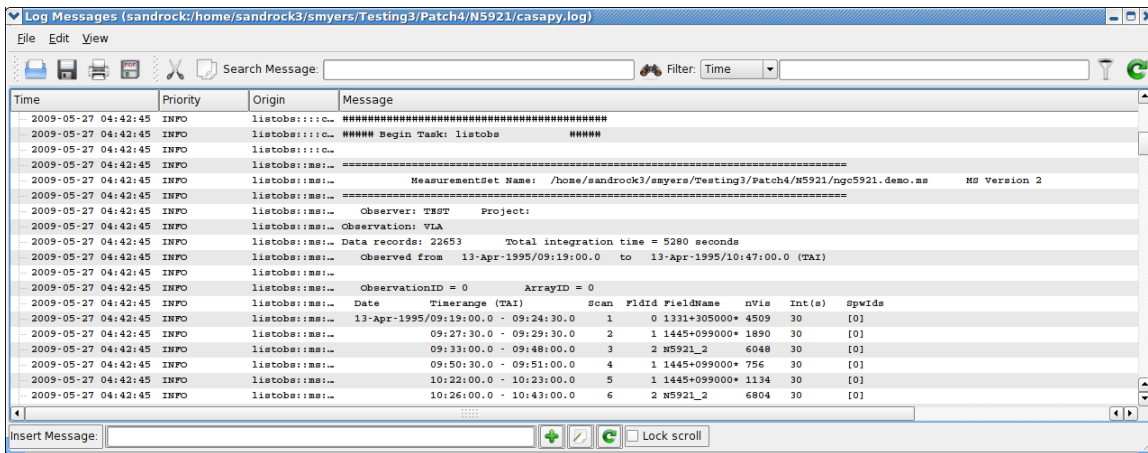


Figure 1.4: The CASA Logger GUI window under Linux. Note that under MacOSX a stripped down logger will instead appear as a Console.

The output contained in `casapy-YYYYMMDD-HHMMSS.log` is also displayed in a separate window using the `casalogger`. Generally, the logger window will be brought up when `casapy` is started. If you do not want the logger GUI to appear, then start `casapy` using the `--nologger` option,

```
casapy --nologger
```

which will run CASA in the terminal window. See § 1.5.2.1 for more startup options.

ALERT: Due to problems with Qt under MacOSX, we had to replace the GUI `qtcasalogger` with a special stripped down one that uses the Mac Console. This still has the important capabilities such as showing the messages and cut/paste. The following description is for the Linux version and thus should mostly be disregarded on OSX. On the Mac, you treat this as just another console window and use the usual mouse and hot-key actions to do what is needed.

The CASA logger window for Linux is shown in Figure 1.4. The main feature is the display area for the log text, which is divided into columns. The columns are:

- **Time** — the time that the message was generated. Note that this will be in local computer time (usually UT) for `casapy` generated messages, and may be different for user generated messages;
- **Priority** — the *Priority Level* (see below) of the message;
- **Origin** — where within CASA the message came from. This is in the format `Task::Tool::Method` (one or more of the fields may be missing depending upon the message);
- **Message** — the actual text.

The `casalogger` GUI has a range of features, which include:

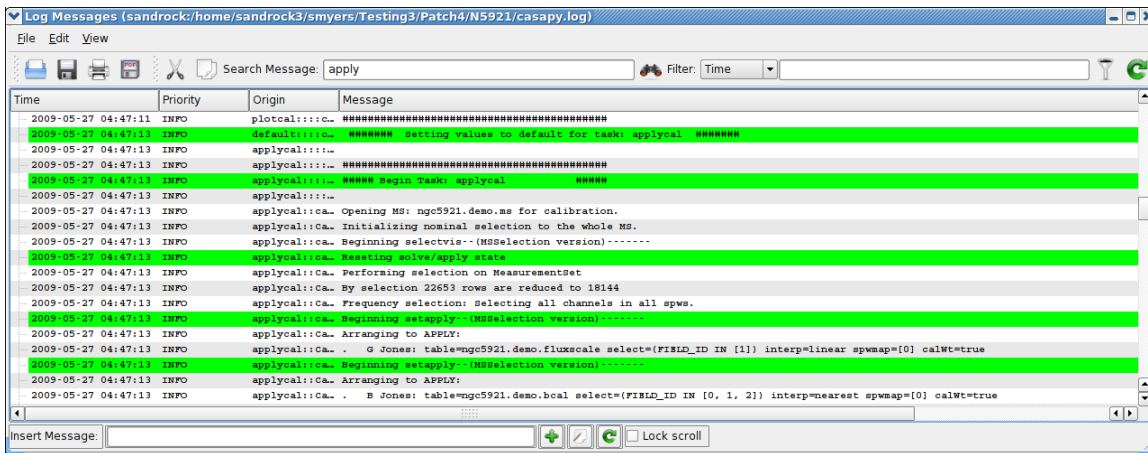


Figure 1.5: Using the Search facility in the `casalogger`. Here we have specified the string 'apply' and it has highlighted all instances in green.

- **Search** — search messages by entering text in the Search window and clicking the search icon. The search currently just matches the exact text you type anywhere in the message. See Figure 1.5 for an example.
- **Filter** — a filter to sort by message priority, time, task/tool of origin, and message contents. Enter text in the Filter window and click the filter icon to the right of the window. Use the pull-down at the left of the Filter window to choose what to filter. The matching is for the exact text currently (no regular expressions). See Figure 1.6 for an example.
- **View** — show and hide columns (Time, Priority, Origin, Message) by checking boxes under the **View** menu pull-down. You can also change the font here.
- **Insert Message** — insert additional comments as “notes” in the log. Enter the text into the “Insert Message” box at the bottom of the logger, and click on the Add (+) button, or choose to enter a longer message. The entered message will appear with a priority of “NOTE” with the Origin as your username. See Figure 1.7 for an example. **ALERT:** This message currently will not be inserted into the correct (or user controllable) order into the log.
- **Copy** — left-click on a row, or click-drag a range of rows, or click at the start and shift click at the end to select. Use the Copy button or **Edit** menu Copy to put the selected rows into the clipboard. You can then (usually) paste this where you wish. **ALERT:** this does not work routinely in the current version. You are best off going to the `casapy-YYYYMMDD-HHMMSS.log` file if you want to grab text.
- **Open** — There is an Open function in the **File** menu, and an Open button, that will allow you to load old `casalogger` files.

Other operations are also possible from the menu or buttons. Mouse “flyover” will reveal the operation of buttons, for example.

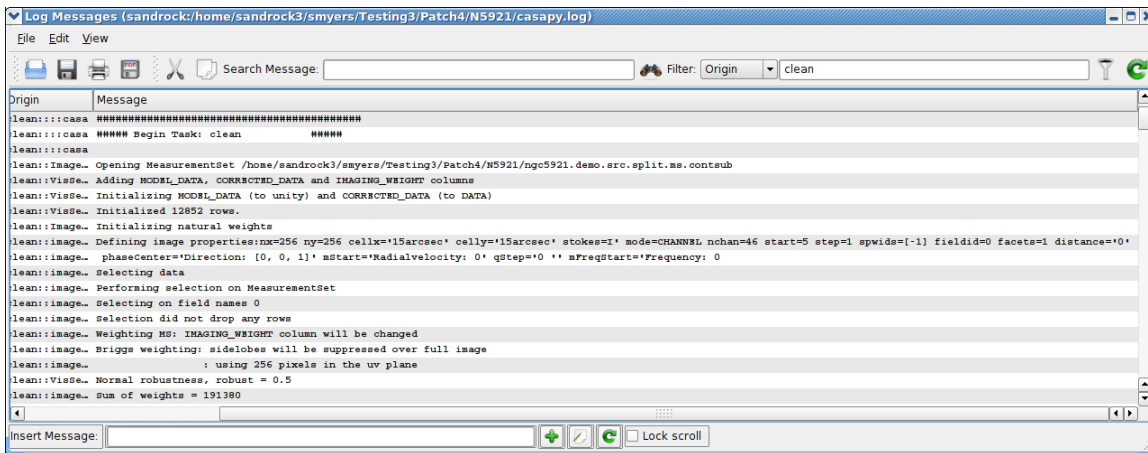


Figure 1.6: Using the `casalogger` Filter facility. The log output can be sorted by Priority, Time, Origin, and Message. In this example we are filtering by Origin using `'clean'`, and it now shows all the log output from the `clean` task.

It is possible to change the file that the logging is directed to. Per default it is `'casapy-YYYYMMDD-HHMMSS.log'`. But starting CASA with the option `--logfile`

```
casapy --logfile otherfile.log
```

will redirect the output of the logger to the file `'otherfile.log'` (see also Sect. 1.5.2.1). The log file can also be changed *during* a CASA session. Type

```
CASA <15>: casalog.setlogfile('otherfile.log')
```

and you will redirect the output to the `'otherfile.log'` file. However, the logger GUI will still be monitoring the previous `'casapy-YYYYMMDD-HHMMSS.log'` file. To change it to the new file, go on **File - Open** and select the new log file, in our case `'otherfile.log'`.

1.5.2.1 Startup options for the logger

One can specify logger options at the startup of `casapy` on the command line:

```
casapy <logger option>
```

The options are described in Appendix A.3. For example, to not bring up a GUI but send the message to your terminal, do

```
casapy --nologger --log2term
```

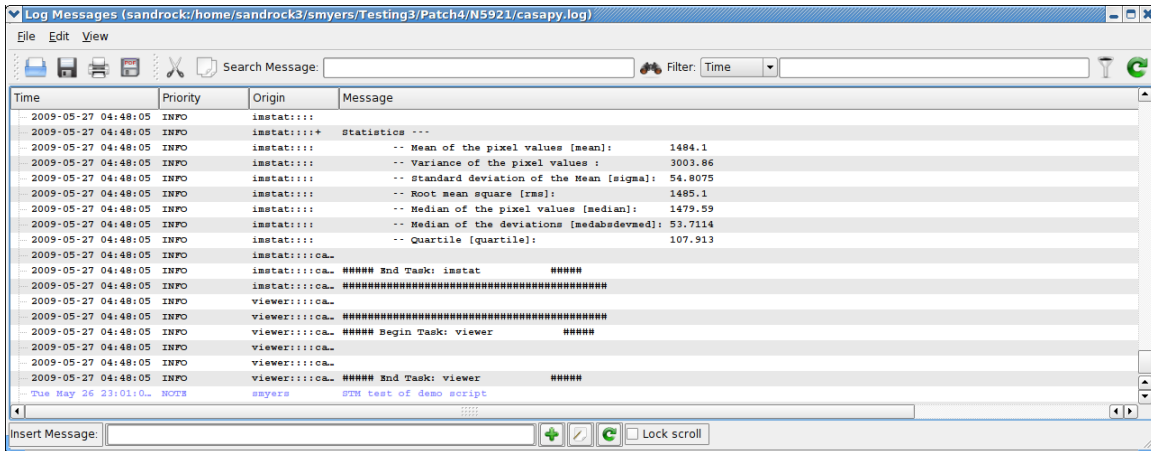


Figure 1.7: CASA Logger - Insert facility: The log output can be augmented by adding notes or comments during the reduction. The file should then be saved to disk to retain these changes.

while

```
casapy --logfile mynewlogfile.log
```

will start casapy with logger messages going to the file mynewlogfile.log.

1.5.2.2 Setting priority levels in the logger

Logger messages are assigned a *Priority Level* when generated within CASA. The current levels of Priority are:

1. **SEVERE** — errors;
2. **WARN** — warnings;
3. **INFO** — basic information every user should be aware of or has requested;
4. **INFO1** — information possibly helpful to the user;
5. **INFO2** — details the power user might want to see;
6. **INFO3** — even more details;
7. **INFO4** — lowest level of non-debugging information;
8. **DEBUGGING** — most “important” debugging messages;
9. **DEBUG1** — more details;

10. **DEBUG2** — lowest level of debugging messages.

The “debugging” levels are intended for the developers use.

There is a threshold for which these messages are written to the `casapy-YYYYMMDD-HHMMSS.log` file and are thus visible in the `logger`. By default, only messages at level **INFO** and above are logged. The user can change the threshold using the `casalog.filter` method. This takes a single string argument of the `level` for the threshold. The `level` sets the lowest priority that will be generated, and all messages of this level or higher will go into the `casapy-YYYYMMDD-HHMMSS.log` file.

Some examples:

```
casalog.filter('INFO')           # the default
casalog.filter('INFO2')         # should satisfy even advanced users
casalog.filter('INFO4')         # all INFOx messages
casalog.filter('DEBUG2')        # all messages including debugging
```

WARNING: Setting the threshold to **DEBUG2** will put lots of messages in the log!

Inside the Toolkit:

The `casalog` tool can be used to control the logging. In particular, the `casalog.filter` method sets the priority threshold. This tool can also be used to change the output log file, and to post messages into the logger.

1.5.3 Where are my data in CASA?

Interferometric data are filled into a so-called Measurement Set (or MS). In its logical structure, the MS looks like a generalized description of data from any interferometric or single dish telescope. Physically, the MS consists of several tables in a directory on disk.

Tables in CASA are actually directories containing files that are the sub-tables. For example, when you create a MS called `AM675.ms`, then the name of the directory where all the tables are stored will be called `AM675.ms/`. See Chapter 2 for more information on Measurement Set and Data Handling in CASA.

The data that you originally get from a telescope can be put in any directory that is convenient to you. Once you “fill” the data into a Measurement Set that can be accessed by CASA, it is generally best to keep that MS in the same directory where you started CASA so you can get access to it easily (rather than constantly having to specify a full path name).

When you generate calibration solutions or images (again these are in table format), these will also be written to disk. It is a good idea to keep them in the directory in which you started CASA.

1.5.3.1 How do I get rid of my data in CASA?

Note that when you delete a Measurement Set, calibration table, or image, which are in fact directories, you must delete this and all underlying directories and files. If you are not running `casapy`, this is most simply done by using the file delete method of the operating system you

started CASA from. For example, when running CASA on a Linux system, in order to delete the Measurement Set named `AM675.ms` type:

```
CASA <5>: !rm -r AM675.ms
```

from within CASA. The `!` tells CASA that a system command follows (see § 1.3.7.5), and the `-r` makes sure that all subdirectories are deleted recursively.

It is convenient to prefix all MS, calibration tables, and output files produced in a run with a common string. For example, one might prefix all files from VLA project AM675 with `AM675`, e.g. `AM675.ms`, `AM675.cal`, `AM675.clean`. Then,

```
CASA <6>: !rm -r AM675*
```

will clean up all of these.

In scripts, the `!` escape to the OS will not work. Instead, use the `os.system()` function (Appendix B.7.1) to do the same thing:

```
os.system('rm -r AM675*')
```

If you are within `casapy`, then the CASA system is keeping a cache of tables that you have been using and using the OS to delete them will confuse things. For example, running a script that contains `rm` commands multiple times will often not run or crash the second time as the cache gets confused. The clean way of removing CASA tables (MS, caltables, images) inside `casapy` is to use the `rmtables` task:

```
rmtables('AM675.ms')
```

and this can also be wildcarded

```
rmtables('AM675*')
```

(though you may get warnings if it tries to delete files or directories that fit the name wildcard that are not CASA tables).

ALERT: Some CASA processes lock the file and forget to give it up when they are done (`plotxy` is usually the culprit). You will get `WARNING` messages from `rmtables` and your script will probably crash second time around as the file isn't removed. The safest thing is still to exit `casapy` and start a new session for multiple runs.

1.5.4 What's in my data?

The actual data is in a large `MAIN` table that is organized in such a way that you can access different parts of the data easily. This table contains a number of “rows”, which are effectively a single timestamp for a single spectral window (like an IF from the VLA) and a single baseline (for an interferometer).

There are a number of “columns” in the MS, the most important of which for our purposes is the `DATA` column — this contains the original visibility data from when the MS was created or filled. There are other helpful “scratch” columns which hold useful versions of the data or weights for further processing: the `CORRECTED_DATA` column, which is used to hold calibrated data and an optional `MODEL_DATA` column, which may hold the Fourier inversion of a particular model image. The creation and use of the scratch columns is generally done behind the scenes, but you should be aware that they are there (and when they are used). We will occasionally refer to the rows and columns in the MS.

More on the contents of the MS can be found in § 2.1.

1.5.5 Data Selection in CASA

We have tried to make the CASA task interface as uniform as possible. If a given parameter appears in multiple tasks, it should, as far as is possible, mean the same thing and be used in the same way in each. There are groups of parameters that appear in a number of tasks to do the same thing, such as for data selection.

The parameters `field`, `spw`, and `selectdata` (which if `True` expands to a number of sub-parameters) are commonly used in tasks to select data on which to work. These common data selection parameters are described in § 2.3.

1.6 From Loading Data to Images

The subsections below provide a brief overview of the steps you will need to load data into CASA and obtain a final, calibrated image. Each subject is covered in more detail in Chapters 2 through 6.

An end-to-end workflow diagram for CASA data reduction for interferometry data is shown in Figure 1.8. This might help you chart your course through the package. In the following subsections, we will chart a rough course through this process, with the later chapters filling in the individual boxes.

Note that single-dish data reduction (for example with the ALMA single-dish system) follows a similar course. This is detailed in Chapter 8.

1.6.1 Loading Data into CASA

The key data and image import tasks are:

- `importuvfits` — import visibility data in UVFITS format (§ 2.2.5);
- `importvla` — import data from VLA that is in *export* format (§ 2.2.3);
- `importasdm` — import ALMA data in ASDM format (§ 2.2.1);

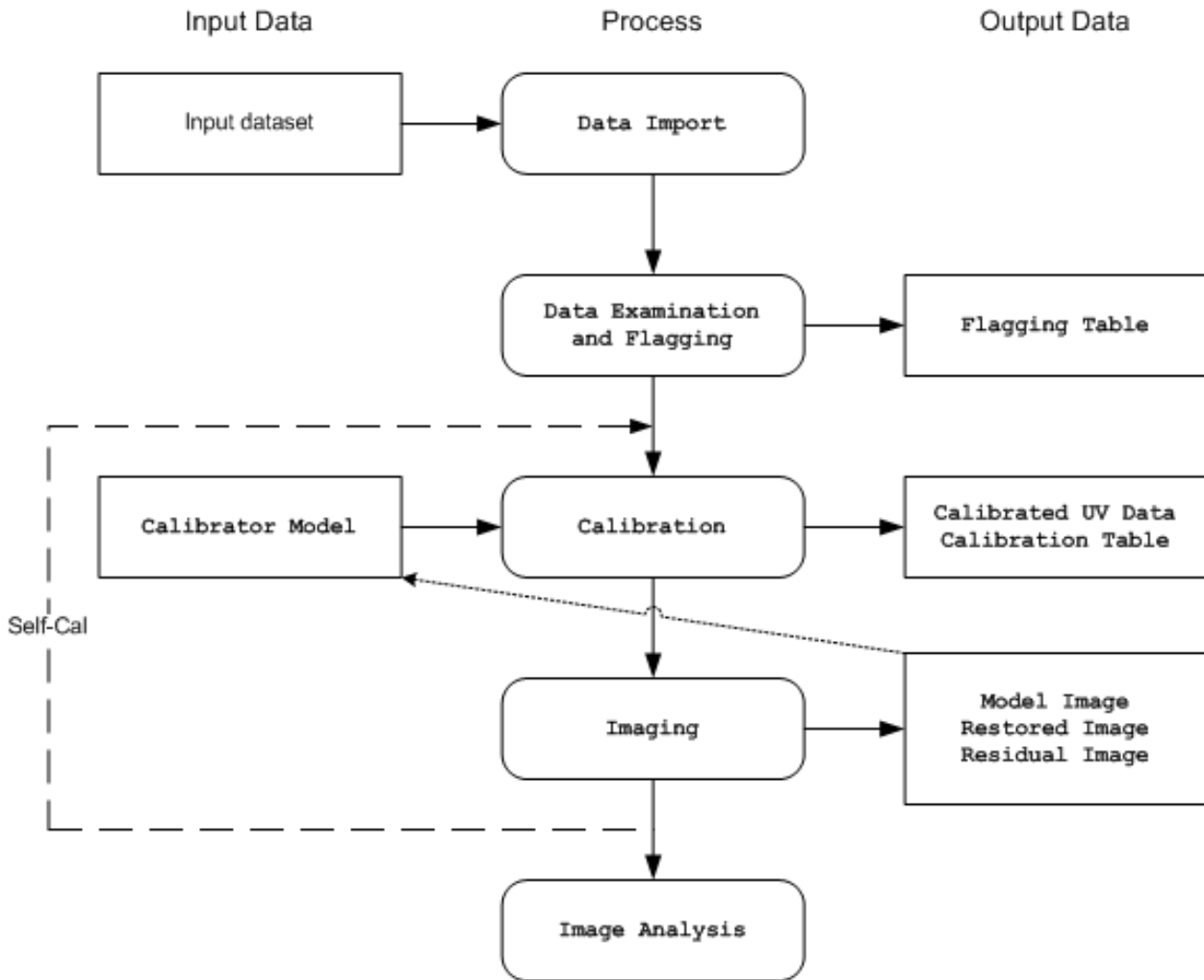


Figure 1.8: Flow chart of the data processing operations that a general user will carry out in an end-to-end CASA reduction session.

- `importevla` — import JVLA/EVLA data in SDM format (§ 2.2.2);
- `importfits` — import a FITS image into a CASA *image* format table (§ 6.24).

These are used to bring in your interferometer data, to be stored as a CASA Measurement Set (MS), and any previously made images or models (to be stored as CASA image tables).

The data import tasks will create a MS with a path and name specified by the `vis` parameter. See § 1.5.3 for more information on MS in CASA. The Measurement Set is the internal data format used by CASA, and conversion from any other native format is necessary for most of the data reduction tasks.

Once data is imported, there are other operations you can use to manipulate the datasets:

- `concat` — concatenate multiple MSs into a given or a new MS (§ 2.2.12)

Data import, export, concatenation, and selection detailed in Chapter 2.

1.6.1.1 VLA: Filling data from VLA archive format

VLA data in “archive” format are read into CASA from disk using the `importvla` task (see § 2.2.3). This filler supports the new naming conventions of EVLA antennas when incorporated into the old VLA system.

Note that future data from the EVLA in ASDM format will use a different filler. This will be made available in a later release.

1.6.1.2 Filling data from UVFITS format

For UVFITS format, use the `importuvfits` task. A subset of popular flavors of UVFITS (in particular UVFITS as written by AIPS) is supported by the CASA filler. See § 2.2.5 for details.

1.6.1.3 Loading FITS images

For FITS format images, such as those to be used as calibration models, use the `importfits` task. Most, though not all, types of FITS images written by astronomical software packages can be read in.

See § 6.24 for more information.

1.6.1.4 Concatenation of multiple MS

Once you have loaded data into Measurement Sets on disk, you can use the tasks `concat` or `virtualconcat` to combine them.

See § 2.2.12 for details.

1.6.2 Data Examination, Editing, and Flagging

The main data examination and flagging tasks are:

- `listobs` — summarize the contents of a MS (§ 2.2.7);
- `flagmanager` — save and manage versions of the flagging entries in the Measurement Set (§ 3.2);
- `plotms` — interactive X-Y plotting and flagging of visibility data (§ 3.3.1);

- **plotxy** — interactive X-Y plotting and flagging of visibility data (§ 3.3.2), **note:** plotxy is slower than plotms and will eventually be phased out, plotxy is still useful to create scripted hardcopy output, this functionality will likely be available in plotms in the next release);
- **flagdata** — flagging (and unflagging) of specified data (§ 3.4);
- **viewer** — the CASA viewer can display (as a raster image) MS data, with some editing capabilities (§ 7);

These tasks allow you to list, plot, and/or flag data in a CASA MS.

There will eventually be tasks for “automatic” flagging to data based upon statistical criteria. Stay tuned.

Examination and editing of synthesis data is described in Chapter 3.

Visualization and editing of an MS using the `casaviewer` is described in Chapter 7.

1.6.2.1 Interactive X-Y Plotting and Flagging

The principal tool for making X-Y plots of visibility data is `plotms` (see § 3.3.1). Amplitudes and phases (among other things) can be plotted against several x-axis options.

Interactive flagging (i.e., “see it – flag it”) is possible on the `plotms` X-Y displays of the data (§ 3.3.1.6). Since flags are inserted into the Measurement Set, it is useful to backup (or make a copy) of the current flags before further flagging is done, using `flagmanager` (§ 3.2). Copies of the flag table can also be restored to the MS in this way.

1.6.2.2 Flag the Data Non-interactively

The `flagdata` task (§ 3.4) will flag the visibility data set based on the specified data selections. The `listobs` task (§ 2.2.7) may be run (e.g. with `verbose=True`) to provide some of the information needed to specify the flagging scope. `flagdata` also contains autoflagging routines.

1.6.2.3 Viewing and Flagging the MS

The CASA `viewer` can be used to display the data in the MS as a (grayscale or color) raster image. The MS can also be edited. Use of the `viewer` on an MS is detailed in § 7.5.

1.6.3 Calibration

The major calibration tasks are:

- **setjy** — Computes the model visibilities for a specified source given a flux density or model image, knows about standard calibrator sources (§ 4.3.5);

- `initweights` — if necessary, supports (re-)initialization of the data weights, including an option for enabling spectral weight accounting (§ 4.3.1)
- `gencal` — Creates a calibration table for known delay and antenna position offsets (§ 4.3.6);
- `bandpass` — Solves for frequency-dependent (bandpass) complex gains (§ 4.4.2);
- `gaincal` — Solves for time-dependent (frequency-independent) complex gains (§ 4.4.3);
- `fluxscale` — Bootstraps the flux density scale from standard calibrators (§ 4.4.4);
- `polcal` — polarization calibration (§ 4.4.5);
- `applycal` — Applies calculated calibration solutions (§ 4.6.1);
- `clearcal` — Re-initializes calibrated visibility data in a given Measurement Set (§ 4.6.3);
- `listcal` — Lists calibration solutions (§ 4.5.2);
- `plotcal` — Plots (and optionally flags) calibration solutions (§ 4.5.1);
- `uvcontsub` — carry out uv-plane continuum subtraction for spectral-line data (§ 4.7.6);
- `split` — write out a new (calibrated) MS for specified sources (§ 4.7.1);
- `cvel` — Regrid a spectral MS onto a new frequency channel system (§ 4.7.7).

During the course of calibration, the user will specify a set of calibrations to pre-apply before solving for a particular type of effect, for example gain or bandpass or polarization. The solutions are stored in a calibration table (subdirectory) which is specified by the user, *not* by the task: care must be taken in naming the table for future use. The user then has the option, as the calibration process proceeds, to accumulate the current state of calibration in a new cumulative table. Finally, the calibration can be applied to the dataset.

Synthesis data calibration is described in detail in Chapter 4.

1.6.3.1 Prior Calibration

The `setjy` task calculates absolute fluxes for Measurement Set base on known calibrator sources. This can then be used in later calibration tasks. Currently, `setjy` knows the flux density as a function of frequency for several standard EVLA flux calibrators and solar system objects, and the value of the flux density can be manually inserted for any other source. If the source is not well-modeled as a point source, then a model image of that source structure can be used (with the total flux density scaled by the values given or calculated above for the flux density). Models are provided for the standard VLA calibrators.

Antenna gain-elevation curves (e.g. for the EVLA antennas) and atmospheric optical depth corrections (applied as an elevation-dependent function) may be pre-applied before solving for the bandpass and gains. CASA v4.1 was the last version where these specialized calibration were supported by explicit parameters in the calibration tasks (`gaincurve` and `opacity`). As of v4.2, these

parameters have been removed, and gain curves and opacity are supported via `gencal`, which will generate standard calibration tables describing these effects, much as other a priori effects (Tsys, switched power, etc.) are supported.

See § 4.3 for more details.

1.6.3.2 Bandpass Calibration

The `bandpass` task calculates a bandpass calibration solution: that is, it solves for gain variations in frequency as well as in time. Since the bandpass (relative gain as a function of frequency) generally varies much more slowly than the changes in overall (mean) gain solved for by `gaincal`, one generally uses a long time scale when solving for the bandpass. The default 'B' solution mode solves for the gains in frequency slots consisting of channels or averages of channels.

A polynomial fit for the solution (solution type 'BPOLY') may be carried out instead of the default frequency-slot based 'B' solutions. This single solution will span (combine) multiple spectral windows.

Bandpass calibration is discussed in detail in § 4.4.2.

If the gains of the system are changing over the time that the bandpass calibrator is observed, then you may need to do an initial gain calibration (see next step).

1.6.3.3 Gain Calibration

The `gaincal` task determines solutions for the time-based complex antenna gains, for each spectral window, from the specified calibration sources. A solution interval may be specified. The default 'G' solution mode solves for antenna-based gains in each polarization in specified time solution intervals. The 'T' solution mode is the same as 'G' except that it solves for a single solution shared by both polarizations.

A spline fit for the solution (solution type 'GSPLINE') may be carried out instead of the default time-slot based 'G' solutions.

See § 4.4.3 for more on gain calibration.

1.6.3.4 Polarization Calibration

The `polcal` task will solve for any unknown polarization leakage and cross-hand phase terms ('D' and 'X' solutions). The 'D' leakage solutions will work on sources with no polarization and sources with known (and supplied, e.g., using `smodel`) polarization. For sources with unknown polarization tracked through a range in parallactic angle on the sky, using poltype 'D+QU', which will first estimate the calibrator polarization for you.

The solution for the unknown cross-hand polarization phase difference 'X' term requires a polarized source with known linear polarization (Q,U).

Frequency-dependent (i.e., per channel) versions of all of these modes are also supported (polypes 'Df', 'Df+QU', and 'Xf').

See § 4.4.5 for more on polarization calibration.

1.6.3.5 Examining Calibration Solutions

The `plotcal` task (§ 4.5.1) will plot the solutions in a calibration table. The `xaxis` choices include time (for `gaincal` solutions) and channel (e.g. for `bandpass` calibration). The `plotcal` interface and plotting surface is similar to that in `plotxy`. Eventually, `plotcal` will allow you to flag and unflag calibration solutions in the same way that data can be edited in `plotxy`.

The `listcal` task (§ 4.5.2) will print out the calibration solutions in a specified table.

1.6.3.6 Bootstrapping Flux Calibration

The `fluxscale` task bootstraps the flux density scale from “primary” standard calibrators to the “secondary” calibration sources. Note that the flux density scale must have been previously established on the “primary” calibrator(s), typically using `setjy`, and of course a calibration table containing valid solutions for all calibrators must be available.

See § 4.4.4 for more.

1.6.3.7 Correcting the Data

The final step in the calibration process, `applycal` may be used to apply several calibration tables (e.g., from `gaincal` or `bandpass`, along with prior calibration tables). The corrections are applied to the `DATA` column of the visibility, writing the `CORRECTED_DATA` column which can then be plotted (e.g. in `plotxy`), `split` out as the `DATA` column of a new MS, or imaged (e.g. using `clean`). Any existing corrected data are overwritten.

See § 4.6.1 for details.

1.6.3.8 Splitting the Data

After a suitable calibration is achieved, it may be desirable to create one or more new Measurement Sets containing the data for selected sources. This can be done using the `split` task (§ 4.7.1).

Further imaging and calibration (e.g. self-calibration) can be carried out on these split Measurement Sets.

1.6.3.9 UV Continuum subtraction

For spectral line data, continuum subtraction can be performed in the image domain (`imcontsub`) or in the uv domain. For the latter, there are two tasks available: `uvcontsub` subtracts polynomial of desired order from each baseline, defined by line-free channels.

1.6.3.10 Transforming the Data to a new frame

If you want to transform your dataset to a different frequency and velocity frame than the one it was observed in, then you can use the `cvel` task (§ 4.7.7). Alternatively, you can do the regridding during the imaging process in `clean` without running `cvel` before.

1.6.4 Synthesis Imaging

The key synthesis imaging tasks are:

- `clean` — Calculates a deconvolved image based on the visibility data, using one of several `clean` algorithms (§ 5.3);
- `feather` — Combines a single dish and synthesis image in the Fourier plane (§ 5.6).

Most of these tasks are used to take calibrated interferometer data, with the possible addition of a single-dish image, and reconstruct a model image of the sky. **Alert:** The `clean` task is now even more powerful and incorporates the functionality of previous specialized tasks such as `mosaic` and `widefield`.

See Chapter 5 for more on synthesis imaging.

1.6.4.1 Cleaning a single-field image or a mosaic

The CLEAN algorithm is the most popular and widely-studied method for reconstructing a model image based on interferometer data. It iteratively removes at each step a fraction of the flux in the brightest pixel in a defined region of the current “dirty” image, and places this in the model image. The `clean` task implements the CLEAN algorithm for single-field data. The user can choose from a number of options for the particular flavor of CLEAN to use.

Often, the first step in imaging is to make a simple gridded Fourier inversion of the calibrated data to make a “dirty” image. This can then be examined to look for the presence of noticeable emission above the noise, and to assess the quality of the calibration by searching for artifacts in the image. This is done using `clean` with `niter=0`.

The `clean` task can jointly deconvolve mosaics as well as single fields, and also has options to do wide-field and wide-band multi-frequency synthesis imaging.

See § 5.3 for an in-depth discussion of the `clean` task.

1.6.4.2 Feathering in a Single-Dish image

If you have a single-dish image of the large-scale emission in the field, this can be “feathered” in to the image obtained from the interferometer data. This is carried out using the `feather` tasks as the weighted sum in the uv-plane of the gridded transforms of these two images. While not as accurate as a true joint reconstruction of an image from the synthesis and single-dish data together, it is sufficient for most purposes.

See § 5.6 for details on the use of the `feather` task.

1.6.5 Self Calibration

Once a calibrated dataset is obtained, and a first deconvolved model image is computed, a “self-calibration” loop can be performed. Effectively, the model (not restored) image is passed back to another calibration process (on the target data). This refines the calibration of the target source, which up to this point has had (usually) only external calibration applied. This process follows the regular calibration procedure outlined above.

Any number of self-calibration loops can be performed. As long as the images are improving, it is usually prudent to continue the self-calibration iterations.

This process is described in § 5.10.

1.6.6 Data and Image Analysis

The key data and image analysis tasks are:

- `imhead` — summarize and manipulate the “header” information in a CASA image (§ 6.2);
- `imcontsub` — perform continuum subtraction on a spectral-line image cube (§ 6.4);
- `immath` — perform mathematical operations on or between images (§ 6.6);
- `immoments` — compute the moments of an image cube (§ 6.7);
- `imstat` — calculate statistics on an image or part of an image (§ 6.9);
- `imval` — extract values of one or more pixels, as a spectrum for cubes, from an image (§ 6.10);
- `imfit` — simple 2D Gaussian fitting of single components to a region of an image (§ 6.5);
- `imregrid` — regrid an image onto the coordinate system of another image (§ 6.13);
- `viewer` — there are useful region statistics and image cube plotting capabilities in the viewer (§ 7).

1.6.6.1 What's in an image?

The `imhead` task will print out a summary of image “header” keywords and values. This task can also be used to retrieve and change the header values.

See § 6.2 for more.

1.6.6.2 Image statistics

The `imstat` task will print image statistics. There are options to restrict this to a box region, and to specified channels and Stokes of the cube. This task will return the statistics in a Python dictionary return variable.

See § 6.9 for more.

1.6.6.3 Image values

The `imval` task will values from an image. There are options to restrict this to a box region, and to return specified channels and Stokes of the cube as a spectrum. This task will return these values in a Python dictionary return variable which can then be operated on in the `casapy` environment.

See § 6.9 for more.

1.6.6.4 Moments of an image cube

The `immoments` task will compute a “moments” image of an input image cube. A number of options are available, from the traditional true moments (zero, first, second) and variations thereof, to other images such as median, minimum, or maximum along the moment axis.

See § 6.7 for details.

1.6.6.5 Image math

The `immath` task will allow you to form a new image by mathematical combinations of other images (or parts of images). This is a powerful, but tricky, task to use.

See § 6.6 for more.

1.6.6.6 Regridding an Image

It is occasionally necessary to regrid an image onto a new coordinate system. The `imregrid` task can be used to regrid an input image onto the coordinate system of an existing template image, creating a new output image.

See § 6.13 for a description of this task.

1.6.6.7 Displaying Images

To display an image use the `viewer` task. The viewer will display images in raster, contour, or vector form. Blinking and movies are available for spectral-line image cubes. To start the viewer, type:

```
viewer
```

Executing the `viewer` task will bring up two windows: a viewer screen showing the data or image, and a file catalog list. Click on an image or ms from the file catalog list, choose the proper display, and the image should pop up on the screen. Clicking on the wrench tool (second from left on upper left) will obtain the data display options. Most functions are self-documenting.

The viewer can be run outside of `casapy` by typing `casaviewer`.

See § 7 for more on viewing images.

1.6.7 Getting data and images out of CASA

The key data and image export tasks are:

- `exportuvfits` — export a CASA MS in UVFITS format (§ 2.2.5);
- `exportfits` — export a CASA image table as FITS (§ 6.24).

These tasks can be used to export a CASA MS or image to UVFITS or FITS respectively. See the individual sections referred to above for more on each.

Chapter 2

Visibility Data Import, Export, and Selection

To use CASA to process your data, you first will need to get it into a form that is understood by the package. These are “Measurement Sets” for synthesis (and single dish) data, and “image tables” for images.

There are a number of tasks used to fill telescope-specific data, to import/export standard formats, to list data contents, and to concatenate multiple datasets. These are:

- `asdmsummary` — list the contents of a archive file in *ASDM* format (§ 2.2.1)
- `importasdm` — import of ALMA data in *ASDM* format (§ 2.2.1)
- `importevla` — import of Jansky VLA data and flags in *ASDM* format (§ 2.2.2)
- `importuvfits` — import visibility data in UVFITS format (§ 2.2.5.1)
- `importfitsidi` — import visibility data in the FITS-IDI format (§ 2.2.5.2)
- `importvla` — import data from VLA that is in *export* format (§ 2.2.3)
- `importmiriad` — import data from MIRIAD visibilities (§ 2.2.4)
- `exportuvfits` — export a CASA MS in UVFITS format (§ 2.2.5.3)
- `listobs` — summarize the contents of a MS (§ 2.2.7)
- `listpartition` — List the summary of a Multi-MS data set in the logger or in a file (§ 2.2.8)
- `listvis` — list the data in a MS (§ 2.2.9)
- `vishead` — list and change the metadata contents of a MS (§ 2.2.10)
- `visstat` — statistics on data in a MS (§ 2.2.11)

- `concat` — concatenate two or more MS into a new MS (§ 2.2.12)
- `virtualconcat` — concatenate two or more MS or multi-MS into a new multi-MS (§ 2.2.12 and chapter 10)

In CASA, there is a standard syntax for selection of data that is employed by multiple tasks. This is described in § 2.3.

There are also tasks for the import and export of image data using FITS:

- `listfits` — show the header content of any FITS file (§ 2.1)
- `importfits` — import a FITS image into a CASA *image* format table (§ 6.24)
- `exportfits` — export a CASA image table as FITS (§ 6.24)

2.1 CASA Measurement Sets

Data is handled in CASA via the `table` system. In particular, visibility data are stored in a CASA table known as a Measurement Set (MS). Details of the physical and logical MS structure are given below, but for our purposes here an MS is just a construct that contains the data. An MS can also store single dish data (essentially a set of auto-correlations of a 1-element interferometer), though there are also data formats more suitable for single-dish spectra (see § 8).

A full description of the Measurement Set can be found at <http://casa.nrao.edu/Memos/229.html>.

Note that images are handled through special `image` tables, although standard FITS I/O is also supported. Images and `image` data are described in a separate chapter.

The headers of any FITS files can be displayed in the logger with the `listfits` task:

```
# listfits :: List the HDU and typical data rows of a fits file:
fitsfile      =      '      # Name of input fits file
```

Unless your data was previously processed by CASA or software based upon its predecessor `aips++`, you will need to import it into CASA as an MS. Supported formats include some “standard” flavors of UVFITS, the VLA “Export” archive format, and most recently, the ALMA Science Data Model (ASDM) format. These are described below in § 2.2.

Once in Measurement Set form, your data can be accessed through various tools and tasks with a common interface. The most important of these is the *data selection interface* (§ 2.3) which allows you to specify the subset of the data on which the tasks and tools will operate.

2.1.1 Under the Hood: Structure of the Measurement Set

Inside the Toolkit:

Measurement sets are handled in the `ms` tool. Import and export methods include `ms.fromfits` and `ms.tofits`.

It is not necessary that a casual CASA user know the specific details on how the data in the MS is stored and the contents of all the sub-tables. However, we will occasionally refer to specific “columns” of the MS when describing the actions of various tasks, and thus we provide the following synopsis to familiarize the user with the necessary nomenclature. You may skip ahead to subsequent sections if you like!

Inside the Toolkit:

Generic CASA tables are handled in the `tb` tool. You have direct access to keywords, rows and columns of the tables with the methods of this tool.

All CASA data files, including Measurement Sets, are written into the current working directory by default, with each CASA table represented as a separate sub-directory. MS names therefore need only comply with UNIX file or directory naming conventions, and can be referred to from within CASA directly, or via full path names.

An MS consists of a **MAIN** table containing the visibility data. and associated sub-tables containing auxiliary or secondary information. The tables are logical constructs, with contents located in the physical `table.*` files on disk. The **MAIN** table consists of the `table.*` files in the main directory of the ms-file itself, and the other tables are in the respective subdirectories. The various MS tables and sub-tables can be seen by listing the contents of the MS directory itself (e.g. using Unix `ls`), or via the `browsetable` task (§ 3.6).

See Fig 2.1 for an example of the contents of a MS directory. Or, from the `casapy` prompt,

```
CASA <1>: ls ngc5921.ms
IPython system call: ls -F ngc5921.ms
ANTENNA          POLARIZATION      table.f1          table.f3_TSM1    table.f8
DATA_DESCRIPTION PROCESSOR          table.f10         table.f4          table.f8_TSM1
FEED             SORTED_TABLE      table.f10_TSM1    table.f5          table.f9
FIELD           SOURCE            table.f11         table.f5_TSM1    table.f9_TSM1
FLAG_CMD        SPECTRAL_WINDOW  table.f11_TSM1    table.f6          table.info
HISTORY         STATE            table.f2          table.f6_TSM0    table.lock
OBSERVATION     table.dat        table.f2_TSM1     table.f7
POINTING        table.f0         table.f3          table.f7_TSM1
```

Note that the **MAIN** table information is contained in the `table.*` files in this directory. Each of the sub-table sub-directories contain their own `table.dat` and other files, e.g.

```
CASA <2>: ls ngc5921.ms/SOURCE
IPython system call: ls -F ngc5921.ms/SOURCE
table.dat table.f0 table.f0i table.info table.lock
```

Each “row” in a table contains entries for a number of specified “columns”. For example, in the **MAIN** table of the MS, the original visibility data is contained in the **DATA** column — each “cell” contains a matrix of observed complex visibilities for that row at a single time stamp, for a single baseline in a single spectral window. The shape of the data matrix is given by the number of channels and the number of correlations (voltage-products) formed by the correlator for an array.

Table 2.1 lists the non-data columns of the **MAIN** table that are most important during a typical data reduction session. Table 2.2 lists the key data columns of the **MAIN** table of an interferometer MS. The MS produced by fillers for specific instruments may insert special columns, such as

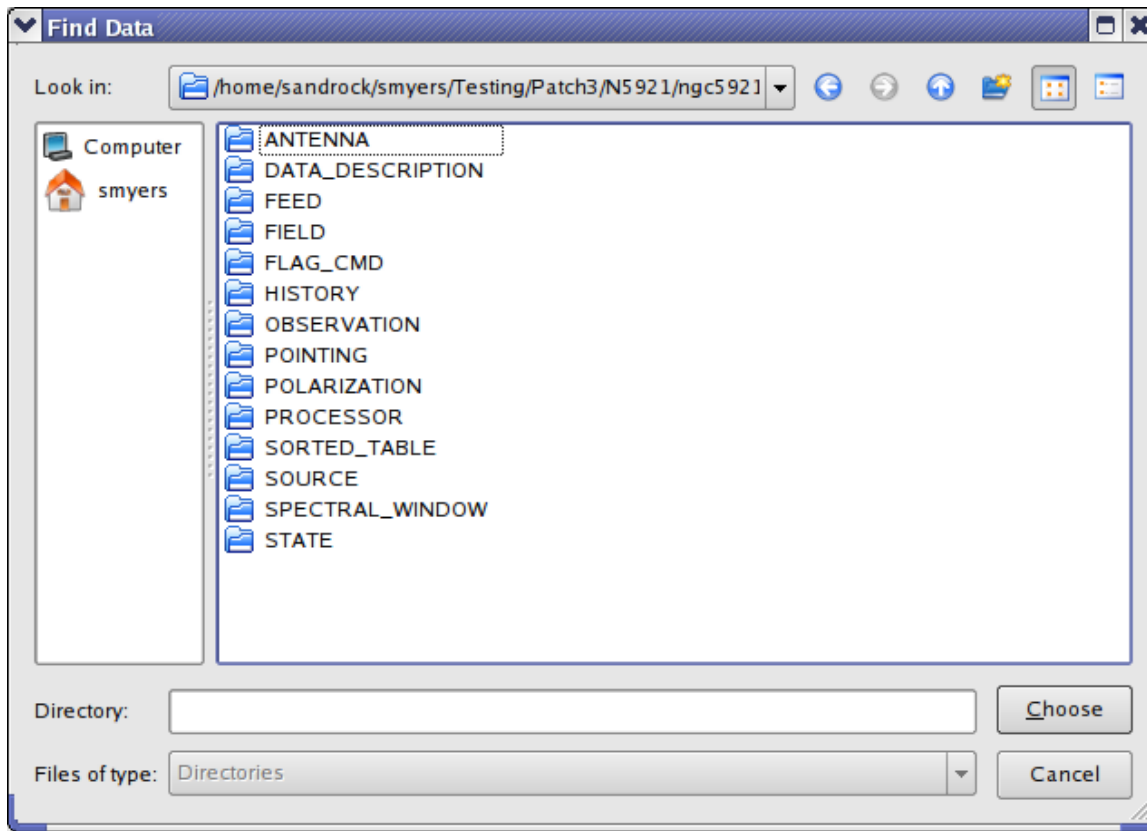


Figure 2.1: The contents of a Measurement Set. These tables compose a Measurement Set named `ngc5921.demo.ms` on disk. This display is obtained by using the **File:Open** menu in `browsetable` and left double-clicking on the `ngc5921.demo.ms` directory.

`ALMA_PHASE_CORR`, `ALMA_NO_PHAS_CORR` and `ALMA_PHAS_CORR_FLAG_ROW` for ALMA data filled using the `importasdm` filler (§ 2.2.1). These columns are visible in `browsetable` and are accessible from the toolkit in the `ms` tool (e.g. the `ms.getdata` method) and from the `tb` “table” tool (e.g. using `tb.getcol`).

Note that when you examine table entries for IDs such as `FIELD_ID` or `DATA_DESC_ID`, you will see 0-based numbers.

The MS can contain a number of “scratch” columns, which are used to hold useful versions of other columns such as the data or weights for further processing. The most common scratch columns are:

- `CORRECTED_DATA` — used to hold calibrated data for imaging or display;
- `MODEL_DATA` — holds the Fourier inversion of a particular model image for calibration or imaging. This column is optional.

Table 2.1: Common columns in the MAIN table of the MS.

Parameter	Contents
ANTENNA1	First antenna in baseline
ANTENNA2	Second antenna in baseline
FIELD_ID	Field (source no.) identification
DATA_DESC_ID	Spectral window number, polarization identifier pair (IF no.)
ARRAY_ID	Subarray number
OBSERVATION_ID	Observation identification
POLARIZATION_ID	Polarization identification
SCAN_NUMBER	Scan number
TIME	Integration midpoint time
UVW	UVW coordinates

The creation and use of the scratch columns is generally done behind the scenes, but you should be aware that they are there (and when they are used).

Table 2.2: Commonly accessed MAIN Table data-related columns. Note that the columns ALMA_PHASE_CORR, ALMA_NO_PHAS_CORR and ALMA_PHAS_CORR_FLAG_ROW are specific to ALMA data filled using the `importasdm` filler.

Column	Format	Contents
DATA	Complex(N_c, N_f)	complex visibility data matrix (= ALMA_PHASE_CORR by default)
FLAG	Bool(N_c, N_f)	cumulative data flags
WEIGHT	Float(N_c)	weight for a row
WEIGHT_SPECTRUM	Float(N_c, N_f)	individual weights for a data matrix
ALMA_PHASE_CORR	Complex(N_c, N_f)	on-line phase corrected data (<i>Not in VLA data</i>)
ALMA_NO_PHAS_CORR	Bool(N_c, N_f)	data that has not been phase corrected (<i>Not in VLA data</i>)
ALMA_PHAS_CORR_FLAG_ROW	Bool(N_c, N_f)	flag to use phase-corrected data or not (<i>not in VLA data</i>)
MODEL_DATA	Complex(N_c, N_f)	Scratch: created by calibrator or imager tools
CORRECTED_DATA	Complex(N_c, N_f)	Scratch: created by calibrator or imager tools

Data flags can be set in the MS, too. Whenever a flag is set, the data will be ignored in all processing steps but not physically deleted from the MS. The flags are channel-based and stored in the MS FLAG subtable. Backups can be stored in the 'MS.flagversions' file that can be accessed

via the `flagmanager` (§ 3.2).

The most recent specification for the MS is **Aips++ Measurement Set definition version 2.0** (<http://casa.nrao.edu/Memos/229.html>).

2.2 Data Import and Export

There are a number of tasks available to bring data in various forms into CASA as a Measurement Set:

- ALMA Science Data Model format data can be imported into CASA (`importasdm`)
- JVLA Science Data Model format data, including online flags, can imported into CASA (`importevla`)
- VLA Archive format data can be imported into CASA (`importvla`)
- UVFITS format can be imported into and exported from CASA (`importuvfits`, `importfitsidi`, and `exportuvfits`)

2.2.1 ALMA: Filling of Science Data Model (ASDM) data

The ALMA and JVLA projects have agreed upon a common archival science data model (ASDM) format, and have jointly developed the software to fill this data into CASA. In the ASDM format, the bulk of the data is contained in large binary data format (BDF) tables, with the meta-data and ancillary information in XML tables. This is structured as a directory, like the MS, and was designed to be similar to the MS to facilitate conversion.

Under the Hood:

The `importasdm` task is just an interface to the stand-alone `asdm2MS` application. To find out the command-line arguments to this application, do `asdm2MS --help`.

The content of an ASDM can be listed with the task `asdmsummary`:

```
# asdmsummary :: Summarized description of an ASDM dataset.
asdm           =           '           # Name of input ASDM directory
```

with an output that contains the list and positions of the antennas, followed by the parameters of each scan like observation time, source name, frequency and polarization setup:

```
Input ASDM dataset : TDEM0008.sb3373760.eb3580330.55661.22790537037
=====
ASDM dataset :TDEM0008.sb3373760.eb3580330.55661.22790537037
=====
```


Exec Block : ExecBlock_0
 Telescope : JVLA
 Configuration name : B
 Observer name : Dr. Juergen Ott
 The exec block started on 2011-04-10T05:28:13.200000000 and ended on 2011-04-10T10:27:12.300000256

27 antennas have been used in this exec block.

Id	Name	Make	Station	Diameter	X	Y	Z
Antenna_0	ea01	UNDEFINED	W36	25	-1606841.96	-5042279.689	3551913.017
Antenna_1	ea02	UNDEFINED	E20	25	-1599340.8	-5043150.965	3554065.219
Antenna_2	ea03	UNDEFINED	E36	25	-1596127.728	-5045193.751	3552652.421
Antenna_3	ea04	UNDEFINED	W28	25	-1604865.649	-5042190.04	3552962.365
Antenna_4	ea05	UNDEFINED	W08	25	-1601614.091	-5042001.653	3554652.509
Antenna_5	ea06	UNDEFINED	N24	25	-1600930.06	-5040316.397	3557330.397
Antenna_6	ea07	UNDEFINED	E32	25	-1597053.116	-5044604.687	3553058.987
Antenna_7	ea08	UNDEFINED	N28	25	-1600863.684	-5039885.318	3557965.319
Antenna_8	ea09	UNDEFINED	E24	25	-1598663.09	-5043581.392	3553767.029
Antenna_9	ea10	UNDEFINED	N32	25	-1600781.039	-5039347.456	3558761.542
Antenna_10	ea11	UNDEFINED	E04	25	-1601068.79	-5042051.91	3554824.835
Antenna_11	ea12	UNDEFINED	E08	25	-1600801.926	-5042219.366	3554706.448
Antenna_12	ea14	UNDEFINED	W12	25	-1602044.903	-5042025.824	3554427.832
Antenna_13	ea15	UNDEFINED	W24	25	-1604008.742	-5042135.828	3553403.707
Antenna_14	ea16	UNDEFINED	N12	25	-1601110.052	-5041488.079	3555597.439
Antenna_15	ea17	UNDEFINED	W32	25	-1605808.656	-5042230.082	3552459.202
Antenna_16	ea18	UNDEFINED	N16	25	-1601061.961	-5041175.88	3556058.022
Antenna_17	ea19	UNDEFINED	W04	25	-1601315.893	-5041985.32	3554808.305
Antenna_18	ea20	UNDEFINED	N36	25	-1600690.606	-5038758.734	3559632.061
Antenna_19	ea21	UNDEFINED	E12	25	-1600416.51	-5042462.45	3554536.041
Antenna_20	ea22	UNDEFINED	N04	25	-1601173.979	-5041902.658	3554987.518
Antenna_21	ea23	UNDEFINED	E16	25	-1599926.104	-5042772.967	3554319.789
Antenna_22	ea24	UNDEFINED	W16	25	-1602592.854	-5042054.997	3554140.7
Antenna_23	ea25	UNDEFINED	N20	25	-1601004.709	-5040802.809	3556610.133
Antenna_24	ea26	UNDEFINED	W20	25	-1603249.685	-5042091.404	3553797.803
Antenna_25	ea27	UNDEFINED	E28	25	-1597899.903	-5044068.676	3553432.445
Antenna_26	ea28	UNDEFINED	N08	25	-1601147.94	-5041733.837	3555235.956

Number of scans in this exec Block : 234

scan #1 from 2011-04-10T05:28:13.200000000 to 2011-04-10T05:33:35.500000256

Intents : OBSERVE_TARGET

Sources : 1331+305=3C286

Subscan #1 from 2011-04-10T05:28:13.200000000 to 2011-04-10T05:33:35.500000256

Intent : UNSPECIFIED

Number of integrations : 322

Binary data in uid:///evla/bdf/1302413292901

Number of integrations : 322

Time sampling : INTEGRATION

Correlation Mode : CROSS_AND_AUTO

Spectral resolution type : FULL_RESOLUTION

Atmospheric phase correction : AP_UNCORRECTED

SpectralWindow_0 : numChan = 256, frame = TOP0,

firstChan = 848400000, chandWidth = 125000 x Polarization_0 : corr = RR,LL

scan #2 from 2011-04-10T05:33:35.500000256 to 2011-04-10T05:35:35.200000000

```

Intents : OBSERVE_TARGET
Sources : 1331+305=3C286
Subscan #1 from 2011-04-10T05:33:35.500000256 to 2011-04-10T05:35:35.200000000
Intent : UNSPECIFIED
Number of integrations : 119

Binary data in uid:///evla/bdf/1302413293280
Number of integrations : 119
Time sampling : INTEGRATION
Correlation Mode : CROSS_AND_AUTO
Spectral resolution type : FULL_RESOLUTION
Atmospheric phase correction : AP_UNCORRECTED
SpectralWindow_0 : numChan = 256, frame = TOPO,
firstChan = 8484000000, chandWidth = 125000 x Polarization_0 : corr = RR,LL

scan #3 from 2011-04-10T05:35:35.200000000 to 2011-04-10T05:36:34.999999488
Intents : OBSERVE_TARGET
Sources : 1331+305=3C286
Subscan #1 from 2011-04-10T05:35:35.200000000 to 2011-04-10T05:36:34.999999488
...

```

The **importasdm** task will fill SDM1.2 and SDM1.3 format data into a CASA visibility data set (MS). ALMA data was recorded in SDM1.2 format from October 2009 until May 2011. Since May 2011, ALMA is using the SDM 1.3 format. In particular all science data from cycle 0 will be in SDM1.3. The JVLA also started using SDM1.2 in October 2009 and continues to use this format as of October 2011. **importasdm** can read all of the above formats. The parameter `useversion` can be used to enable the options `process_syspower`, `process_caldevice`, and `process_pointing`.

The default inputs of **importasdm** are:

```

# importasdm :: Convert an ALMA Science Data Model observation into a
CASA visibility file (MS) or single-dish data format (Scantable)
asdm           =      ''      # Name of input asdm directory (on
                             # disk)
vis            =      ''      # Root name of the ms to be created.
                             # Note the .ms is NOT added
createms       =      False    # Create a multi-MS output
singledish     =      False    # Set true to output single-dish data
                             # format
corr_mode      =      'all'    # specifies the correlation mode to be
                             # considered on input. A quoted string
                             # containing a sequence of ao, co,
                             # ac, or all separated by whitespaces
                             # is expected
srt            =      'all'    # specifies the spectral resolution
                             # type to be considered on input. A
                             # quoted string containing a sequence
                             # of fr, ca, bw, or all separated by
                             # whitespaces is expected
time_sampling  =      'all'    # specifies the time sampling
                             # (INTEGRATION and/or SUBINTEGRATION)

```

```

# to be considered on input. A quoted
# string containing a sequence of i,
# si, or all separated by whitespaces
# is expected
ocorr_mode      =      'ca'      # output data for correlation mode
# AUTO_ONLY (ao) or CROSS_ONLY (co) or
# CROSS_AND_AUTO (ca)
compression     =      False     # Flag for turning on data compression
lazy            =      False     # Make the MS DATA column read the ASDM
# Binary data directly (faster import,
# smaller MS)
asis            =      ''        # Creates verbatim copies of the
# ASDMtables in the ouput measurement
# set. Value given must be a string
# of table names separated by spaces;
# A * wildcard is allowed.
wvr_corrected_data =      'no'    # Specifies which values are considered
# in the SDM binary data to fill the
# DATA column in the MAIN table of the
# MS. Expected values for this option
# are: no, for uncorrected data
# (default), yes, for the corrected
# data, and both, for for corrected
# and uncorrected data. Note if both
# is selected two measurement sets are
# created, one with uncorrected data
# and the other with corrected data.
scans           =      ''        # processes only the specified scans.
# This value is a semicolon separated
# list of scan specifications. A scan
# specification consists in an exec
# block index followed by the :
# character; followed by a comma
# separated list of scan indexes or
# scan index ranges. A scan index is
# relative to the exec block it
# belongs to. Scan indexes are 1-based
# while exec blocks are 0-based. "0:1"
# or "2:2~6" or
# "0:1,1:2~6,8;2:,3:24~30" "1,2" are
# valid values for the option. "3:"
# alone will be interpreted as, all
# the scans of the exec block#3. An
# scan index or a scan index range not
# preceded by an exec block index will
# be interpreted as, all the scans
# with such indexes in all the exec
# blocks. By default all the scans
# are considered.
ignore_time     =      False     # All the rows of the tables Feed,
# History, Pointing, Source, SysCal,

```

```

# CalDevice, SysPower, and Weather are
# processed independently of the time
# range of the selected exec block /
# scan.
process_syspower      =      True      # The SysPower table is processed if
# and only if this parameter is set to
# true.
process_caldevice     =      True      # The CalDevice table is processed if
# and only if this parameter is set to
# true.
process_pointing      =      True      # The Pointing table is processed if
# and only if this parameter is set to
# true. If set to False, the POINTING
# table is empty in the resulting MS
process_flags         =      True      # Create online flags in the FLAG_CMD
# sub-table.
    tbuff              =      0.0      # Time padding buffer (seconds)
    applyflags         =      False     # Apply the flags to the MS.
    savecmds          =      False     # Save flag commands to an ASCII file
    outfile            =      ''       # Name of ASCII file to save flag
# commands

flagbackup            =      True      # Back up flag column before applying
# flags.
verbose               =      False     # Output lots of information while the
# filler is working
overwrite             =      False     # Over write an existing MS(s)
showversion           =      False     # Report the version of asdm2MS being
# used
useversion            =      'v3'     # Version of asdm2MS to be used ('v3'
# (default, should work for all data))
bdfflags              =      False     # Set the MS FLAG column according to
# the ASDM_binary_flags
with_pointing_correction =      False  # add (ASDM::Pointing::encoder -
# ASDM::Pointing::pointingDirection)
# to the value to be written in
# MS::Pointing::direction
remove_ref_undef     =      False     # if set to True then apply
# fixspwbackport on the resulting
# MS(es).
convert_ephem2geo    =      True      # if True, convert any attached
# ephemerides to the GEO reference
# frame (time-spacing not changed)

```

If `scans` is set, then `importasdm` processes only the scans specified in the option's value. This value is a semicolon separated list of scan specifications. A scan specification consists in an exec block index followed by the character `':'` followed by a comma separated list of scan indexes or scan index ranges. A scan index is relative to the exec block it belongs to. Scan indexes are 1-based while exec blocks are 0-based. The expressions

```
"0:1"
"2:2~6"
"0:1,1:2~6,8;2:,3:24~30"
"1,2"
"3:"
```

are all valid values for the selection. The "3:" selector will be interpreted as *'all the scans of the exec block 3'*. An scan index or a scan index range not preceded by an exec block index will be interpreted as *'all the scans with such indexes in all the exec blocks'*. By default all the scans are considered.

When `process_flags=True` the task will create online flags based on the `Flag.xml`, `Antenna.xml` and `SpectralWindow.xml` files and copy them to the `FLAG_CMD` sub-table of the MS. The flags will *NOT* be applied unless the parameter `applyflags` is set to `True`. Optionally, the flags can also be saved to an external ASCII file if `savecmds` is set to `True`.

When `bdfflags=True` the task will apply online flags contained in the ASDM BDF data by calling the executable `bdfflags2MS` which the user can also do from the OS prompt. This is recommended for ALMA data.

If `singledish=True`, output data format is scantable (single-dish data format, see 8) instead of MS. In that case, you must specify name or id of the antenna that you want to obtain data. This can be done by using antenna parameter that is defined as a subparameter of `singledish`. For single-dish mode, only auto-correlation data are filled, i.e. `ocorr_mode` is forcibly set to 'ao'.

The option `createms` prepares the output file for parallel processing and creates a multi-MS (see Sect. 10, and 10.2).

2.2.1.1 Import of ASDM data with option `lazy=True`

With release 4.3, the parameter *'lazy'* (default = *False*) is fully tested and operational. If the default value *False* is chosen, `importasdm` will (as in previous versions) fill the visibilities into a newly created `DATA` column of the MS converting them from their binary format in the ASDM to the CASA Table format.

If, however, *lazy* is set to *True*, the task will create the `DATA` column with an ALMA data-specific storage manager, the (`asdmstman`), which enables CASA to *directly* read the binary data from the ASDM with on-the-fly conversion. *No redundant copy of the raw data is created.*

This procedure has the advantage that it saves more than 60% disk space and at least in some cases makes the access to the `DATA` column $\geq 10\%$ faster because the data I/O volume is decreased. For the same reason, it also accelerates the import itself by ca. a factor 2. The acceleration is particularly large in the `applycal` task and here particularly on standard SATA disks.

E.g., if your ASDM has a size of 36 GB, the import with default parameters will turn this into an MS of 73 GB size (total disk space consumption = 36 GB + 73 GB = 109 GB). With `lazy=True`, the imported MS has a size of only 2 GB (total disk space consumption = 36 GB + 2 GB = 38 GB). I.e. your total disk space savings are ca. 65%. Even when you compare to the case where you

delete the ASDM after normal import, the solution with lazy import and keeping the ASDM will save you ca. 48% disk space (in the example above 38 GB compared to 73 GB).

The only caveats are the following:

1. You must not delete your ASDM. You can, however, move it but you have to update the reference stored in the MS. Symbolic links will work. See below on how to use the tool method `ms.asdmref()` to manipulate the ASDM reference.
2. The *lazily* imported DATA column is read-only. But in any normal data reduction, the DATA column (as opposed to CORRECTED_DATA) is treated as read-only anyway.

The *lazily* imported MS is numerically identical with the traditionally imported MS and so are all results derived from the MSs. The setting `lazy=True` might be made the default setting in future CASA releases.

An important additional tool to manipulate lazily imported MSs is the new method `ms.asdmref()` in the `ms` tool. If the MS is imported from an ASDM with option `lazy=True`, the DATA column of the MS is virtual and directly reads the visibilities from the ASDM. A reference to the original ASDM is stored with the MS. If the ASDM needs to be moved to a different path, the reference to it in the MS needs to be updated. This can be achieved with `ms.asdmref()`.

The method takes one argument: `abspath`. When called with `abspath` equal to an empty string (default), the method just reports the currently set ASDM path or an empty string if the ASDM path was not set, i.e. the MS was not lazily imported.

If you want to move the referenced ASDM to a different path, you can set the new *absolute* path by providing it as the value of `abspath` to the method.

```
ms.open('uid___A12345_X678_X910.ms', False)
ms.asdmref('/home/alma/myanalysis/uid___A12345_X678_X910')
ms.close()
```

will set the new location of the referenced ASDM to `/home/alma/myanalysis/uid___A12345_X678_X910`.

Note that the lazily imported *MS* can be moved without any restrictions independently from the referenced ASDM as long as the absolute path to the ASDM remains accessible, even across file systems.

2.2.2 Jansky VLA: Filling of Science Data Model (ASDM) data

The `importevla` task will fill SDM data from the Jansky VLA (or ALMA) into a MS, along with online flagging data contained in the `Flag.xml` SDM table. Otherwise, it behaves as `importasdm` but with a streamlined parameter set.

The default inputs are:

Under the Hood:

The `importevla` task is a modified version of the `importasdm` task, that includes import of online flags from the `Flag.xml` table into the `FLAG_CMD` MS table, and a streamlined set of parameters.

```

# importevla :: Convert an Science Data Model observation into a CASA Measurement Set
asdm          =      ''      # Name of input asdm directory (on disk)
vis           =      ''      # Root name of the ms to be created. Note the .ms
                                # is NOT added
ocorr_mode    =      'co'    # Fill correlation mode AUTO_ONLY (ao),
                                # CROSS_ONLY (co) or CROSS_AND_AUTO (ca)
compression   =      False   # Flag for turning on data compression
asis         =      ''      # Create verbatim copies of these SDM tables in
                                # the MS.
scans         =      ''      # List of scans to fill (default is all scans).
verbose       =      False   # Output lots of information while the filler is
                                # working
overwrite     =      False   # Over write an existing MS
online        =      True    # Create online flags
    tbuff     =      0.0    # Time padding buffer (in seconds)

flagzero      =      True    # Create flag commands for zero points
    flagpol   =      True    # Create flag commands for cross-hand
                                # correlations

shadow        =      True    # Create flag commands for shadowed data
    tolerance =      0.0    # Amount of shadow allowed (in meters)
    addantenna =      ''    # File name or dictionary with additional antenna
                                # names, positions and diameters

applyflags    =      False   # Apply flag commands to MS
savecmds      =      False   # Save flag commands to an ASCII file
flagbackup    =      True    # Back up flag column before applying flags

```

ALERT: If you want to use your JVLA online flags then you must use `importevla` rather than `importasdm`. The `flagcmd` task will process these flags. Also, if you have run `importevla` in CASA 3.3 or earlier, the flag syntax will be processed by the task `oldflagcmd`.

The default action of `importevla` is to construct the `FLAG_CMD` MS table based on the settings of `online`, `flagzero`, and `shadow` (and sub-parameters). If `applyflags=True` then these flags will be applied after filling. We recommend you use the `flagcmd` task after filling to examine these flags and then apply.

See `importasdm` (§ 2.2.1) for a description of the common parameters. Some differences:

Note that `importevla` automatically loads in VLA switched power information (unlike in previous versions).

The `online` parameter controls creation of online flags from the `Flag.xml` SDM table. The `tbuff` parameter adds a time “buffer” padding for these flags in both directions to deal with timing mismatches. **ALERT:** For JVLA data taken before April 2011, you should set `tbuff` to a value (in seconds) equal to $1.5 \times$ the integration time.

The `flagzero` parameter controls creation of clipping commands to flag visibilities with amplitudes that are exact zeros. If `flagpol=True` then it will flag the cross-hands (e.g. RL and LR) as well,

which might result in low but correct values of these correlations being thrown out (but can catch erroneous zeros also). **ALERT:** This facility is provided as the JVLA correlator, particularly in 2010, occasionally produces visibilities with zero or very small values that need to get flagged out.

The `shadow` parameter turns on creation of flag commands to remove antenna time ranges where they are shadowed by other antennas in the array. By default it will flag based on the antenna diameter, but if you want more lenient or conservative flagging then set the `tolerance` sub-parameter, where the shadowed antennas are flagged for all baselines that are shorter than $\text{radius}_1 + \text{radius}_2 - \text{tolerance}$ (the radii are those for the antennas as listed in the ANTENNA sub-table). `addantenna` can be a file that defines the positions of antennas that are on the ground but do not appear in the MS. They can still shadow antennas in the array.

`savecmds` will save all flagging commands in the `flagdata` and `flagcmd` syntax (§ 3.4 and 3.5) to a file to be applied later or for bookkeeping.

A flag backup can be performed using the `flagbackup` parameter. It saves all current flags to the `*.flagversions` file of the MS, before all new flags are applied.

2.2.3 VLA: Filling data from archive format (`importvla`)

VLA data in archive format (i.e., as downloaded from the VLA data archive) are read into CASA from disk using the `importvla` task. The inputs are:

```
# importvla :: import VLA archive file(s) to a measurement set:

archivefiles =      '' # Name of input VLA archive file(s)
vis           =      '' # Name of output visibility file
bandname      =      '' # VLA frequency band name:''=>obtain all bands in archive files
frequencytol = 150000.0 # Frequency shift to define a unique spectral window (Hz)
project       =      '' # Project name: '' => all projects in file
starttime    =      '' # start time to search for data
stoptime     =      '' # end time to search for data
applytsys    =      True # apply nominal sensitivity scaling to data & weights
autocorr     =      False # import autocorrelations to ms, if set to True
antnamescheme = 'new' # 'old' or 'new'; 'VA04' or '4' for ant 4
keepblanks   =      False # Fill scans with empty source names (e.g. tipping scans)?
evlabands    =      False # Use updated eVLA frequencies and bandwidths
```

The main parameters are `archivefiles` to specify the input VLA Archive format file names, and `vis` to specify the output MS name.

ALERT: The scaling of VLA data both before and after the June 2007 Modcomp-turnoff is fully supported, based on the value of `applytsys`.

The NRAO Archive is located at:

- <https://archive.nrao.edu>

Note that `archivefiles` takes a string or list of strings, as there are often multiple files for a project in the archive.

For example:

```
archivefiles = ['AP314_A950519.xp1', 'AP314_A950519.xp2']
vis = 'NGC7538.ms'
```

The `importvla` task allows selection on the frequency band. Suppose that you have 1.3 cm line observations in K-band and you have copied the archive data files `AP314_A950519.xp*` to your working directory and started `casapy`. Then,

```
default('importvla')
archivefiles = ['AP314_A950519.xp1', 'AP314_A950519.xp2', 'AP314_A950519.xp3']
vis = 'ngc7538.ms'
bandname = 'K'
frequencytol = 10e6
importvla()
```

If the data is located in a different directory on disk, then use the full path name to specify each archive file, e.g.:

```
archivefiles=['/home/rohir2/jmcmulli/ALMATST1/Data/N7538/AP314_A950519.xp1', \
              '/home/rohir2/jmcmulli/ALMATST1/Data/N7538/AP314_A950519.xp2', \
              '/home/rohir2/jmcmulli/ALMATST1/Data/N7538/AP314_A950519.xp3']
```

Important Note: `importvla` will import the on-line flags (from the VLA system) along with the data. Shadowed antennas will also be flagged. The flags will be put in the `MAIN` table and thus available to subsequent tasks and tools. If you wish to revert to unflagged data, use `flagmanager` (§ 3.2) to save the flags (if you wish), and then use `flagdata` (§ 3.4) with `mode='manualflag'` and `unflag=True` to toggle off the flags.

The other parameters are:

2.2.3.1 Parameter `applytsys`

The `applytsys` parameter controls whether the nominal sensitivity scaling (based on the measured `TSYS`, with the weights scaled accordingly using the integration time) is applied to the visibility amplitudes or not. If `True`, then it will be scaled so as to be the same as `AIPS FILLM` (i.e. approximately in deciJanskys). Note that post-Modcomp data is in raw correlation coefficient and will be scaled using the `TSYS` values, while Modcomp-era data had this applied online. In all cases `importvla` will do the correct thing to data and weights based on an internal flag in the VLA Archive file, either scaling it or unscaling based on your choice for `applytsys`.

If `applytsys=True` and you see strange behavior in data amplitudes, it may be due to erroneous `TSYS` values from the online system. You might want to then fill with `applytsys=False` and look at the correlation coefficients to see if the behavior is as expected.

2.2.3.2 Parameter bandname

The `bandname` indicates the VLA Frequency band(s) to load, using the traditional bandname codes. These are:

- '4' = 48-96 MHz
- 'P' = 298-345 MHz
- 'L' = 1.15-1.75 GHz
- 'C' = 4.2-5.1 GHz
- 'X' = 6.8-9.6 GHz
- 'U' = 13.5-16.3 GHz
- 'K' = 20.8-25.8 GHz
- 'Q' = 38-51 GHz
- '' = all bands (default)

Note that as the transition from the VLA to JVLA progresses, the actual frequency ranges covered by the bands will expand, and additional bands will be added (namely 'S' from 1-2 GHz and 'A' from 26.4-40 GHz).

2.2.3.3 Parameter frequencytol

The `frequencytol` parameter specifies the frequency separation tolerated when assigning data to spectral windows. The default is `frequencytol=150000` (Hz). For Doppler tracked data, where the sky frequency changes with time, a `frequencytol < 10000` Hz may produce too many unnecessary spectral windows.

2.2.3.4 Parameter project

You can specify a specific `project` name to import from archive files. The default '' will import data from all projects in file(s) `archivefiles`.

For example for VLA Project AL519:

```
project = 'AL519'    # this will work
project = 'al519'    # this will also work
```

while `project='AL0519'` will NOT work (even though that is what queries to the VLA Archive will print it as - sorry!).

2.2.3.5 Parameters starttime and stoptime

You can specify start and stop times for the data, e.g.:

```
starttime = '1970/1/31/00:00:00'
stoptime = '2199/1/31/23:59:59'
```

Note that the blank defaults will load all data fitting other criteria.

2.2.3.6 Parameter autocorr

Note that autocorrelations are filled into the data set if `autocorr=True`. Generally for the VLA, autocorrelation data is not useful, and furthermore the imaging routine will try to image the autocorrelation data (it assumes it is single dish data) which will swamp any real signal. Thus, if you do fill the autocorrelations, you will have to flag them before imaging.

2.2.3.7 Parameter antnamescheme

The `antnamescheme` parameter controls whether `importvla` will try to use a naming scheme where JVLA antennas are prefixed with `EA` (e.g. `'EA16'`) and old VLA antennas have names prefixed with `VA` (e.g. `'VA11'`). Our method to detect whether an antenna is JVLA is not yet perfected, and thus unless you require this feature, simply use `antnamescheme='old'`.

2.2.3.8 Parameter evlabands

The `evlabands=True` option is provided to allow users to access JVLA frequencies outside the standard VLA tunings (e.g. the extended C-band above 6 GHz). **ALERT:** use of this option for standard VLA data will cause unexpected associations, such as X-band data below 8 GHz being extracted to C-band (as the JVLA C-band is 4–8 GHz). Use with care.

2.2.4 Import MIRIAD visibilities (`importmiriad`)

The task `importmiriad` allows one to import visibilities in the MIRIAD data format to be converted to a MS. The task has mainly been tested on data from the ATCA and CARMA telescopes and the inputs are:

```
# importmiriad :: Convert a Miriad visibility file into a CASA MeasurementSet
mirfile      =      ''      # Name of input Miriad visibility file
vis          =      ''      # Name of output MeasurementSet
tsys         =      False   # Use the Tsys to set the visibility weights
spw          =      'all'   # Select spectral windows
vel          =      ''      # Select velocity reference (TOP0,LSRK,LSRD)
linecal      =      False   # (CARMA) Apply line calibration
wide         =      'all'   # (CARMA) Select wide window averages
debug        =      0       # Display increasingly verbose debug messages
```

For details, please contact ATCA or CARMA staff.

2.2.5 UVFITS Import and Export

The UVFITS format is not exactly a standard, but is a popular archive and transport format nonetheless. CASA supports UVFITS files written by the AIPS FITTP task, and others.

UVFITS is supported for both import and export.

2.2.5.1 Import using `importuvfits`

To import UVFITS format data into CASA, use the `importuvfits` task:

```
CASA <1>: inp(importuvfits)
fitsfile      =      '' # Name of input UVFITS file
vis           =      '' # Name of output visibility file (MS)
antnamescheme =      'old' # For VLA only; 'new' or 'old'; 'VA04' or '04' for VLA ant 4
```

This is straightforward, since all it does is read in a UVFITS file and convert it as best it can into a MS.

For example:

```
importuvfits(fitsfile='NGC5921.fits',vis='ngc5921.ms')
```

ALERT: CARMA data can be loaded into CASA. However,

```
tb.open("c0104I/ANTENNA",nomodify=False)
namelist=tb.getcol("NAME").tolist()
for i in range(len(namelist)):
name = 'CA'+namelist[i]
print ' Changing '+namelist[i]+' to '+name
namelist[i]=name

tb.putcol("NAME",namelist)
tb.close()
```

2.2.5.2 Import using `importfitsidi`

Some uvfits data is written in the FITS-IDI standard. Those files can be imported into CASA with the `importfitsidi` task:

```
# importfitsidi :: Convert a FITS-IDI file to a CASA visibility data set
fitsidifile      =      ['']      # Name(s) of input FITS-IDI file(s)
vis              = 'ngc5921.demo.ms' # Name of output visibility file (MS)
constobsid      =      False      # If True, give constant obs ID==0 to
                                # the data from all input fitsidi
                                # files (False = separate obs id for
                                # each file)
scanreindexgap_s =      0.0        # min time gap (seconds) between
                                # integrations to start a new scan
```

The *constobs* parameter can be used to give all visibilities the same observation id of 0. *scanreindexgap_s* controls the gap that defines different scans.

Example:

```
importuvfits(fitsidifile='NGC1300.fits',vis='NGC1300.ms')
```

2.2.5.3 Export using exportuvfits

The *exportuvfits* task will take a MS and write it out in UVFITS format. The defaults are:

```
# exportuvfits :: Convert a CASA visibility data set (MS) to a UVFITS file

vis          =      ''      # Name of input visibility file
fitsfile     =      ''      # Name of output UVFITS file)
datacolumn   = 'corrected' # which data to write (data, corrected, model)
field        =      ''      # Field name list
spw          =      ''      # Spectral window and channel selection
antenna      =      ''      # antenna list to select
time         =      ''      # time range selection
nchan        =      -1     # Number of channels to select
start        =      0      # Start channel
width        =      1      # Channel averaging width (value>1 indicates averaging)
writesyscal  =      False  # Write GC and TY tables
multisource  =      True   # Write in multi-source format
combinespw   =      True   # Combine spectral windows (True for AIPS)
writestation =      True   # Write station name instead of antenna name
```

For example:

```
exportuvfits(vis='ngc5921.split.ms',
             fitsfile='NGC5921.split.fits',
             multisource=False)
```

The MS selection parameters *field*, *spw*, *antenna*, and *timerange* follow the standard selection syntax described in § 2.3.

ALERT: The *nchan*, *start*, and *width* parameters will be superseded by channel selection in *spw*. Currently, there is a *time* parameter rather than *timerange*.

The `datacolumn` parameter chooses which data-containing column of the MS (see § 2.1.1) is to be written out to the UV FITS file. Choices are: `'data'`, `'corrected'`, and `'model'`.

There are a number of special parameters that control what is written out. These are mostly here for compatibility with AIPS.

The `writesyscal` parameter toggles whether `GC` and `TY` extension tables are written. These are important for VLBA data, and for JVLA data. **ALERT:** Not yet available.

The `multisource` parameter determines whether the UV FITS file is a multi-source file or a single-source file, if you have a *single-source* MS or choose only a single source. Note: the difference between a single-source and multi-source UVFITS file here is whether it has a source (SU) table and the source ID in the random parameters. Some programs (i.e. difmap) only accept single-source files. If you select more than one source in fields, then the multisource parameter will be overridden to be `True` regardless.

The `combinespw` parameter allows, if some conditions are met, exporting all of spectral windows (SpW) as a set of "IF"s in a single "FREQID" setup instead of giving each SpW its own FREQID in the FITS file. In this context an IF (Intermediate Frequency) is a specialization of an SpW, where each IF in a UV FITS file must have the same number of channels and polarizations, each channel must have the same width, and each IF must be present (even if flagged) throughout the entire observation. If these conditions are not met the data must be exported using multiple FREQIDs, the UV FITS equivalent of a general SpW. This matters since many (sub)programs will work with multiple IFs, but not multiple FREQIDs. For example, a UV FITS file with multiple FREQIDs can be read by AIPS, but you may find that you have to separate the FREQIDs with `SPLIT` before you can do very much with them. Therefore `combinespw=True` should be `True` if possible. Typically MSes where each band was observed simultaneously can be exported with `combinespw=True`. MSes where the tuning changed with time, e.g. 10 minutes at 4.8 GHz followed by 15 minutes at 8.4 GHz, should be exported to multiple UV FITS files using `spw` to select one tuning (set of simultaneous SpWs) per file.

The `multisource` parameter determines whether the UV FITS file is a multi-source file or a single-source file, if you have a *single-source* MS or choose only a single source. Note: the difference between a single-source and multi-source UVFITS file here is whether it has a source (SU) table and the source ID in the random parameters. If you select more than one source in fields, then the multisource parameter will be overridden to be `True` regardless.

The `combinespw` parameter allows combination of all spectral windows at one time. If `True`, then all spectral windows must have the same shape. For AIPS to read an exported file, then set `combinespw=True`.

The `writestation` parameter toggles the writing of the station name instead of antenna name.

2.2.6 Handling Measurement Set metadata and data

There are tasks provided for basic listing and manipulation of Measurement Set data and metadata. These include:

- `listobs` — summarize the contents of a MS (§ 2.2.7)
- `vishead` — list and change the metadata contents of a MS (§ 2.2.10)
- `visstat` — statistics on data in a MS (§ 2.2.11)
- `concat` — concatenate two or more MS into a new MS (§ 2.2.12)

2.2.7 Summarizing your MS (`listobs`)

Once you import your data into a CASA Measurement Set, you can get a summary of the MS contents with the `listobs` task.

The inputs are:

```
# listobs :: List the summary of a data set in the logger or in a file
vis                = 'day2_TDEM0003_10s_norx' # Name of input visibility file (MS)
selectdata        = True                    # Data selection parameters
  field           = ''                      # Field names or field index
                                                # numbers: '' ==>all, field='0~2,3C286'
  spw             = ''                      # spectral-window/frequency/channel
  antenna         = ''                      # antenna/baselines: ''==>all, antenna ='3,VA04'
  timerange       = ''                      # time range: ''==>all,timerange='09:14:0~09:54:0'
  correlation      = ''                      # Select data based on correlation
  scan            = ''                      # scan numbers: ''==>all
  intent          = ''                      # Select data based on observation intent: ''==>all
  feed            = ''                      # multi-feed numbers: Not yet implemented
  array           = ''                      # (sub)array numbers: ''==>all
  uvrange         = ''                      # uv range: ''==>all; uvrange
                                                # ='0~100klambda', default units=meters
  observation     = ''                      # Select data based on observation ID: ''==>all

verbose           = True
listfile          = ''                      # Name of disk file to write output: ''==>to terminal
listunfl         = False                   # List unflagged row counts?
                                                # If true, it can have significant negative performance impact.
```

The summary (of the selected data) will be written to the logger, to the `casapy-YYYYMMDD-HHMMSS.log` file, and optionally to a file specified in the `listfile` parameter. For example,

```
listobs('n5921.ms')
```

results in the logger messages:

```
listobs(vis="day2_TDEM0003_10s_norx",selectdata=True,spw="",field="",
        antenna="",uvrange="",timerange="",correlation="",scan="",
        intent="",feed="",array="",observation="",verbose=True,
```

```

listfile="",listunfl=False)
=====
MeasurementSet Name: /Users/jott/casa/casatest/casa4.0/irc/day2_TDEM0003_10s_norx      MS Version 2
=====
Observer: Mark J. Mark Claussen      Project: T.B.D.
Observation: EVLA
Data records: 290218      Total integration time = 10016 seconds
Observed from 26-Apr-2010/03:21:56.0 to 26-Apr-2010/06:08:52.0 (UTC)

ObservationID = 0      ArrayID = 0
Date      Timerange (UTC)      Scan  FldId  FieldName      nRows  SpwIds  Average Interval(s)
26-Apr-2010/03:21:51.0 - 03:23:21.0      5      2  J0954+1743      2720  [0, 1]  [10, 10]
           03:23:39.0 - 03:28:25.0      6      3  IRC+10216      9918  [0, 1]  [10, 10]
           03:28:38.0 - 03:29:54.0      7      2  J0954+1743      2700  [0, 1]  [10, 10]
           03:30:08.0 - 03:34:53.5      8      3  IRC+10216      9918  [0, 1]  [10, 10]
...
(nRows = Total number of rows per scan)
Fields: 4
ID  Code Name      RA      Decl      Epoch  SrcId  nRows
2   D   J0954+1743      09:54:56.823626 +17.43.31.22243 J2000  2      65326
3   NONE IRC+10216      09:47:57.382000 +13.16.40.65999 J2000  3      208242
5   F   J1229+0203      12:29:06.699729 +02.03.08.59820 J2000  5      10836
7   E   J1331+3030      13:31:08.287984 +30.30.32.95886 J2000  7      5814
Spectral Windows: (2 unique spectral windows and 1 unique polarization setups)
SpwID  Name      #Chans  Frame  Ch1(MHz)  ChanWid(kHz)  TotBW(kHz)  Corrs
0      Subband:0  64      TOPO   36387.229  125.000      8000.0      RR  RL  LR  LL
1      Subband:0  64      TOPO   36304.542  125.000      8000.0      RR  RL  LR  LL
Sources: 10
ID  Name      SpwId  RestFreq(MHz)  SysVel(km/s)
0   J1008+0730  0      0.03639232     -0.026
0   J1008+0730  1      0.03639232     -0.026
2   J0954+1743  0      0.03639232     -0.026
2   J0954+1743  1      0.03639232     -0.026
3   IRC+10216  0      0.03639232     -0.026
3   IRC+10216  1      0.03639232     -0.026
5   J1229+0203  0      0.03639232     -0.026
5   J1229+0203  1      0.03639232     -0.026
7   J1331+3030  0      0.03639232     -0.026
7   J1331+3030  1      0.03639232     -0.026
Antennas: 19:
ID  Name  Station  Diam.  Long.      Lat.      Offset from array center (m)      ITRF G
           East      North      Elevation
0   ea01  W09      25.0 m  -107.37.25.2 +33.53.51.0 -521.9407 -332.7782 -1.1977 -1601710.01700
1   ea02  E02      25.0 m  -107.37.04.4 +33.54.01.1 9.8247 -20.4292 -2.7808 -1601150.05950
2   ea03  E09      25.0 m  -107.36.45.1 +33.53.53.6 506.0591 -251.8666 -3.5832 -1600715.94800
3   ea04  W01      25.0 m  -107.37.05.9 +33.54.00.5 -27.3562 -41.3030 -2.7418 -1601189.03014
4   ea05  W08      25.0 m  -107.37.21.6 +33.53.53.0 -432.1158 -272.1493 -1.5032 -1601614.09100
5   ea07  N06      25.0 m  -107.37.06.9 +33.54.10.3 -54.0667 263.8720 -4.2292 -1601162.59320
6   ea08  N01      25.0 m  -107.37.06.0 +33.54.01.8 -30.8810 -1.4664 -2.8597 -1601185.63494
7   ea09  E06      25.0 m  -107.36.55.6 +33.53.57.7 236.9058 -126.3369 -2.4443 -1600951.58800
8   ea12  E08      25.0 m  -107.36.48.9 +33.53.55.1 407.8394 -206.0057 -3.2252 -1600801.91600
9   ea15  W06      25.0 m  -107.37.15.6 +33.53.56.4 -275.8288 -166.7451 -2.0590 -1601447.19800
10  ea19  W04      25.0 m  -107.37.10.8 +33.53.59.1 -152.8599 -83.8054 -2.4614 -1601315.89300
11  ea20  N05      25.0 m  -107.37.06.7 +33.54.08.0 -47.8454 192.6015 -3.8723 -1601168.78610
12  ea21  E01      25.0 m  -107.37.05.7 +33.53.59.2 -23.8638 -81.1510 -2.5851 -1601192.46780

```


13	ea22	N04	25.0 m	-107.37.06.5	+33.54.06.1	-42.5986	132.8623	-3.5431	-1601173.95370
14	ea23	E07	25.0 m	-107.36.52.4	+33.53.56.5	318.0523	-164.1848	-2.6960	-1600880.57000
15	ea24	W05	25.0 m	-107.37.13.0	+33.53.57.8	-210.0944	-122.3885	-2.2581	-1601377.00800
16	ea25	N02	25.0 m	-107.37.06.2	+33.54.03.5	-35.6245	53.1806	-3.1345	-1601180.86148
17	ea27	E03	25.0 m	-107.37.02.8	+33.54.00.5	50.6647	-39.4832	-2.7249	-1601114.36550
18	ea28	N08	25.0 m	-107.37.07.5	+33.54.15.8	-68.9057	433.1889	-5.0602	-1601147.94040

using the (default) `verbose=True` option. The most useful extra information that `verbose=True` gives is the list of the scans in the dataset.

2.2.8 MMS summary (listpartition)

Similar to `listobs`, `listpartition` shows the summary of a Multi-Measurement Set (MMS).

The inputs are:

```
# listpartition :: List the summary of a multi-MS data set in the logger or in a file
vis              =          ''      # Name of multi-MS or normal MS.
createdict       =      False      # Create and return a dictionary with
                                   # sub-MS information
listfile         =          ''      # Name of ASCII file to save output:
                                   # ''==>to terminal
```

For example,

```
listpartition('n5921.mms')
```

results in the logger messages:

```
This is a multi-MS with separation axis = scan,spw
Sub-MS          Scan  Spw   Nchan  Nrows  Size
ngc5921.mms.0000.ms  2   [0]   [63]   1890   27M
                  4   [0]   [63]    756
                  5   [0]   [63]   1134
                  6   [0]   [63]   6804
ngc5921.mms.0001.ms  1   [0]   [63]   4509   28M
                  3   [0]   [63]   6048
                  7   [0]   [63]   1512
```

2.2.9 Listing MS data (listvis)

The `listvis` task will print to the terminal (or file) listing of the data in your MS. The inputs are:

```
# listvis :: List measurement set visibilities.
vis          =      ''      # Name of input visibility file
options     =      'ap'     # List options: ap only
datacolumn  =      'data'   # Column to list: data, float_data, corrected, model,
                             # residual
field       =      ''      # Field names or index to be listed: ''==>all
spw         =      '*'      # Spectral window:channels: '\*'==>all, spw='1:5~57'
selectdata  =      False    # Other data selection parameters
observation  =      ''      # Select by observation ID(s)
average     =      ''      # Averaging mode: ==>none (Not yet implemented)
showflags   =      False    # Show flagged data (Not yet implemented)
pagerows   =      50       # Rows per page
listfile    =      ''      # Output file
```

For example:

Units of columns are: Date/Time(Yymmdd/HH:MM:SS UT), UVDist(wavelength), Phase(deg), UVW(m)

WEIGHT: 7

FIELD: 2

SPW: 0

Date/Time:

Date/Time:			RR:				RL:				LR:				LL:				
2010/04/26/	Intrf	UVDist	Chn	Amp	Phs	Wt	F	Amp	Phs	Wt	F	Amp	Phs	Wt	F	Amp	Phs	Wt	F
03:21:56.0	ea01-ea02	72363	0:	0.005	-124.5	7		0.005	25.7	7		0.001	104.6	7		0.000	23.4	7	
03:21:56.0	ea01-ea02	72363	1:	0.001	-4.7	7		0.001	-135.1	7		0.004	-14.6	7		0.001	19.9	7	
03:21:56.0	ea01-ea02	72363	2:	0.002	17.8	7		0.002	34.3	7		0.005	-114.3	7		0.005	-149.7	7	
03:21:56.0	ea01-ea02	72363	3:	0.004	-19.4	7		0.003	-79.2	7		0.002	-89.0	7		0.004	31.3	7	
03:21:56.0	ea01-ea02	72363	4:	0.001	-16.8	7		0.004	-141.5	7		0.005	114.9	7		0.006	105.2	7	
03:21:56.0	ea01-ea02	72363	5:	0.001	-29.8	7		0.009	-96.4	7		0.002	-125.0	7		0.002	-64.5	7	

...

Type Q to quit, A to toggle long/short list, or RETURN to continue [continue]:

ALERT: We are working on improving the format of the listvis output.

2.2.10 Listing and manipulating MS metadata (vishead)

The vishead task is provided to access keyword information in the Measurement Set. The default inputs are:

```
# vishead :: List, get, and put metadata in a measurement set
vis          =      ''      # Name of input visibility file
mode         =      'list'   # options: list, summary, get, put
listitems    =      []      # items to list ([] for all)
```

The mode = 'summary' option just gives the same output as listobs.

For mode = 'list' the options are: 'telescope', 'observer', 'project', 'field', 'freq_group_name', 'spw_name', 'schedule', 'schedule_type', 'release_date'.

```

CASA <29>: vishead('ngc5921.demo.ms',mode='list',listitems=[])
Out[29]:
{'cal_grp': (array([-1, -1, -1], dtype=int32), {}),
 'field': (array(['1331+30500002_0', '1445+09900002_0', 'N5921_2'],
                dtype='|S16'),
           {}),
 'fld_code': (array(['C', 'A', ''],
                    dtype='|S2'), {}),
 'freq_group_name': (array(['none'],
                            dtype='|S5'), {}),
 'log': ({'r1': False}, {}),
 'observer': (array(['TEST'],
                    dtype='|S5'), {}),
 'project': (array([''],
                   dtype='|S1'), {}),
 'ptcs': ({'r1': array([[[-2.74392758]],
                        [[ 0.53248521]]]),
           'r2': array([[[-2.42044692]],
                        [[ 0.17412604]]]),
           'r3': array([[[-2.26020138]],
                        [[ 0.08843002]]])}),
 {'MEASINFO': {'Ref': 'J2000', 'type': 'direction'},
  'QuantumUnits': array(['rad', 'rad'],
                        dtype='|S4')}),
 'release_date': (array([ 4.30444800e+09]),
                  {'MEASINFO': {'Ref': 'TAI', 'type': 'epoch'},
                   'QuantumUnits': array(['s'],
                                           dtype='|S2')}),
 'schedule': ({'r1': False}, {}),
 'schedule_type': (array([''],
                          dtype='|S1'), {}),
 'source_name': (array(['1331+30500002_0', '1445+09900002_0', 'N5921_2'],
                       dtype='|S16'),
                 {}),
 'spw_name': (array(['none'],
                    dtype='|S5'), {}),
 'telescope': (array(['VLA'],
                     dtype='|S4'), {})}

```

You can use `mode='get'` to retrieve the values of specific keywords, and likewise `mode='put'` to change them. The inputs are:

```

mode      =      'get'      # options: list, summary, get, put
hdkey     =      ''        # keyword to get/put
hdindex   =      ''        # keyword index to get/put, counting from zero. ==>all

```

and

```
# vishead :: List, summary, get, and put metadata in a measurement set
mode      =      'put'    # options: list, summary, get, put
hdkey     =      ''      # keyword to get/put
hdindex   =      ''      # keyword index to get/put, counting from zero. ==>all
hdvalue   =      ''      # value of hdkey
```

For example, a common operation is to change the Telescope name (e.g. if it is unrecognized), e.g.

```
CASA <36>: vishead('ngc5921.demo.ms',mode='get',hdkey='telescope')
```

```
Out [36]:
(array(['VLA'],
      dtype='|S4'), {})
```

```
CASA <37>: vishead('ngc5921.demo.ms',mode='put',hdkey='telescope',hdvalue='JVLA')
```

```
CASA <38>: vishead('ngc5921.demo.ms',mode='get',hdkey='telescope')
```

```
Out [38]:
(array(['JVLA'],
      dtype='|S5'), {})
```

2.2.11 MS statistics (visstat)

ALERT: This is still a prototype task.

The `visstat` task is provided to obtain simple statistics for a Measurement Set, useful in regression tests.

The inputs are:

```
# visstat :: Displays statistical information from a measurement set
vis      =      ''      # Name of input visibility file
axis     =      'amp'   # Which values to use
datacolumn = 'data'   # Which data column to use (data, corrected, model)

useflags =      True   # Take flagging into account?
spw      =      ''     # spectral-window/frequency/channel
field    =      ''     # Field names or field index numbers: ''==>all, field='0~2,3C286'
selectdata = True    # More data selection parameters (antenna, timerange etc)
antenna  =      ''     # antenna/baselines: ''==>all, antenna = '3,VA04'
timerange = ''      # time range: ''==>all, timerange='09:14:0~09:54:0'
correlation = ''    # Select data based on correlation
scan     =      ''     # scan numbers: ''==>all
array    =      ''     # (sub)array numbers: ''==>all
uvrange  =      ''     # uv range: ''==>all; uvrange = '0~100klambda', default units=meters
```

Running this task returns a record (Python dictionary) with the statistics, which can be captured in a Python variable. For example,

```
CASA <42>: mystat = visstat('ngc5921.demo.ms',axis='amp',datacolumn='corrected',field='0')
```

```
CASA <43>: mystat
```

```
  Out[43]:
{'CORRECTED': {'max': 51.938671112060547,
               'mean': 14.796444141750133,
               'medabsdevmed': 0.28020858764648438,
               'median': 14.764373779296875,
               'min': 0.81362706422805786,
               'npts': 514916.0,
               'quartile': 0.56053066253662109,
               'rms': 14.829294204711914,
               'stddev': 0.98650836609147285,
               'sum': 7618925.8316934109,
               'sumsq': 113234125.12642419,
               'var': 0.97319875636846753}}
```

```
CASA <44>: print mystat['CORRECTED']['stddev']
0.986508366091
```

The options for `axis` are:

```
axis='amplitude'    # or ('amp')
axis='phase'
axis='imag'
axis='scan_number'
axis='flag'
```

The phase of a complex number is in radians with range $(-\pi, \pi)$.

2.2.12 Concatenating multiple datasets (concat)

Once you have your data in the form of CASA Measurement Sets, you can go ahead and process your data using the editing, calibration, and imaging tasks. In some cases, you will most efficiently operate on single MS for a particular session (such as calibration). Other tasks will (eventually) take multiple Measurement Sets as input. For others, it is easiest to combine your multiple data files into one.

If you need to combine multiple datasets, you can use the `concat` task. The default inputs are:

```
# concat :: Concatenate several visibility data sets.
vis          =      ''      # Name of input visibility files to be
                        # concatenated
concatvis    =      ''      # Name of output visibility file
freqtol      =      ''      # Frequency shift tolerance for considering data
                        # as the same spwid
dirtol       =      ''      # Direction shift tolerance for considering data
                        # as the same field
```

```

respectname      =      False      # If true, fields with a different name are not
                                     # merged even if their direction agrees
timesort         =      False      # If true, sort by TIME in ascending order
copypointing    =      True       # Copy all rows of the POINTING table.
visweightscale  =      []         # List of the weight scaling factors to be
                                     # applied to the individual MSs

```

The `vis` parameter will take a list of one or more MS. Usually, this will contain all the MS to combine. `concat` will presort the visibilities in time.

With `visweightscale`, a list of weights can be manually specified for the respective input data sets. They will be applied at the time of the combination. To determine the appropriate weights for this procedure, one can inspect the weights (*Wt* and *WtSp* axis parameters) of the input datasets in `plotms`.

The `concatvis` parameter contains the name of the output MS. If this points to an existing file on disk, then the MS in `vis` will be appended to it, otherwise a new MS file is created to contain the concatenated data. Be careful here!

The `timesort` parameter can be used to make sure the output MS is in time order (e.g. if your input MS have concurrent times). This can possibly speed up some subsequent calibration operations.

Furthermore, the parameter `copypointing` can be used to control whether the POINTING table will be carried along in the concatenation process or if the output MS should not contain a POINTING table. This table is quite large for some data (e.g. ALMA) and is mainly needed for mosaic imaging. If you are certain that you will not need it, you can save time and disk space by setting `copypointing` to `False`. Also, as stated in the ephemeris handling section 4.7.11.3, `concat` will correctly merge fields which use the same ephemeris.

The parameters `freqtol` and `dirtol` control how close together in frequency and angle on the sky spectral windows or field locations need to be before calling them the same.

ALERT: Note that if multiple frequencies or pointings are combined using `freqtol` or `dirtol`, then the data are not changed (i.e. not rephased to the single phase center). Use of these parameters is intended to be tolerant of small offsets (e.g. planets tracked which move slightly in J2000 over the course of observations, or combining epochs observed with slightly different positions).

For example:

```

default('concat')
vis = ['n4826_16apr.split.ms', 'n4826_22apr.split.ms']
concatvis = 'n4826_tboth.ms'
freqtol = '50MHz'
visweightscale=['1', '2']
concat()

```

combines the two days in `'n4826_16apr.split.ms'` and `'n4826_22apr.split.ms'` into a new output MS called `'n4826_tboth.ms'`, and the second MS is weighted twice the first one.

ALERT: Note that if you are concatenating MSs which use antennas which were moved between observations, the MS definition does only foresee a unique antenna ID, but not a unique name(!).

The moved antenna will appear twice in the antenna list under the same name but on different *stations* and with two different IDs. The pair ('NAME@STATION') will be the unique identifier.

If you would like to only concatenate the subtables of several MSs, not the bulk visibility data, you can use the task `testconcat` instead of `concat` to save time and disk space. `testconcat` has the same parameters as `concat`. It produces an output MS with the concatenated subtables and an empty Main table.

Furthermore, the task `virtualconcat` permits to concatenate MSs into a multi-MS (MMS, see chapter 10) which is much faster as the data is *moved* into the MMS rather than copied and only some reindexing is done. The bulk data is not rewritten. If you want to keep a copy of the original MSs, set the parameter `keepcopy` of `virtualconcat` to `True`. The creation of that copy will of course consume some of the time you saved by doing a virtual concatenation. Otherwise `virtualconcat` offers the same functionality as `concat`.

2.3 Data Selection

Once in MS form, subsets of the data can be operated on using the tasks and tools. In CASA, there are three common data selection parameters used in the various tasks: `field`, `spw`, and `selectdata`. In addition, the `selectdata` parameter, if set to `True`, will open up a number of other sub-parameters for selection. The selection operation is unified across all the tasks. The available `selectdata` parameters may not be the same in all tasks. But if present, the same parameters mean the same thing and behave in the same manner when used in any task.

For example:

```
field      =      ''      # field names or index of calibrators ''==>all
spw        =      ''      # spectral window:channels: ''==>all
selectdata =      False   # Other data selection parameters
```

versus

```
field      =      ''      # field names or index of calibrators ''==>all
spw        =      ''      # spectral window:channels: ''==>all
selectdata =      True    # Other data selection parameters
  timerange =      ''      # time range: ''==>all
  uvrange   =      ''      # uv range''=all
  antenna   =      ''      # antenna/baselines: ''==>all
  scan      =      ''      # scan numbers: Not yet implemented
  msselect  =      ''      # Optional data selection (Specialized. but see help)
```

The following are the general syntax rules and descriptions of the individual selection parameters of particular interest for the tasks:

2.3.1 General selection syntax

Most of the selections are effected through the use of selection strings. This sub-section describes the general rules used in constructing and parsing these strings. Note that some selections are done through the use of numbers or lists. There are also parameter-specific rules that are described under each parameter.

All lists of basic selection specification-units are comma separated lists and can be of any length. White-spaces before and after the commas (e.g. '3C286, 3C48, 3C84') are ignored, while white-space within sub-strings is treated as part of the sub-string (e.g. '3C286, VIRGO A, 3C84'). In some cases, spaces need to be quoted, e.g. "'spw 1'" (note the double quote around the single quotes).

All integers can be of any length (in terms of characters) composed of the characters 0–9. Floating point numbers can be in the standard format (DIGIT.DIGIT, DIGIT., or .DIGIT) or in the mantissa-exponent format (e.g. 1.4e9). Places where only integers make sense (e.g. IDs), if a floating point number is given, only the integer part is used (it is truncated).

Range of numbers (integers or real numbers) can be given in the format 'N0~N1'. For integer ranges, it is expanded into a list of integers starting from N0 (inclusive) to N1 (inclusive). For real numbers, it is used to select all values present for the appropriate parameter in the Measurement Set between N0 and N1 (including the boundaries). Note that the '~' character is used rather than the more obvious '-' in order to accommodate hyphens in strings and minus signs in numbers.

Wherever appropriate, units can be specified. The units are used to convert the values given to the units used in the Measurement Set. For ranges, the unit is specified only once (at the end) and applies to both the range boundaries.

2.3.1.1 String Matching

String matching can be done in three ways. Any component of a comma separated list that cannot be parsed as a number, a number range, or a physical quantity is treated as a regular expression or a literal string. If the string does not contain the characters '*', '{', '}' or '?', it is treated as a literal string and used for exact matching. If any of the above mentioned characters are part of the string, they are used as a regular expression. As a result, for most cases, the user does not need to supply any special delimiters for literal strings and/or regular expressions. For example:

```
field = '3'      # match field ID 3 and not select field named "3C286".

field = '3*'     # used as a pattern and matched against field names. If
                 # names like "3C84", "3C286", "3020+2207" are found,
                 # all will match. Field ID 3 will not be selected
                 # (unless of course one of the above mentioned field
                 # names also correspond to field ID 3!).

field = '30*'    # will match only with "3020+2207" in above set.
```

However if it is required that the string be matched exclusively as a regular expression, it can be supplied within a pair of '/' as delimiters (e.g. '/.+BAND.+/'). A string enclosed within double

quotes (‘’’’) is used exclusively for pattern matching (patterns are a simplified form of regular expressions - used in most UNIX commands for string matching). Patterns are internally converted to equivalent regular expressions before matching. See the Unix command "info regex", or visit <http://www.regular-expressions.info>, for details of regular expressions and patterns.

Strings can include any character except the following:

```
’,’ ’;’ ’”’ ’/’ NEWLINE
```

(since these are part of the selection syntax). Strings that do not contain any of the characters used to construct regular expressions or patterns are used for exact matches. Although it is highly discouraged to have name in the MS containing the above mentioned reserved characters, if one *does* choose to include the reserved characters as parts of names etc., those names can only be matched against quoted strings (since regular expression and patterns are a super-set of literal strings – i.e., a literal string is also a valid regular expression).

This leaves ’’, ’*’, ’{’, ’}’ or ’?’ as the list of printable character that cannot be part of a name (i.e., a name containing this character can never be matched in a MSSelection expression). These will be treated as pattern-matching even inside double double quotes (’" ”’). There is currently no escape mechanism (e.g. via a backslash).

Some examples of strings, regular expressions, and patterns:

- The string ’LBAND’ will be used as a literal string for exact match. It will match only the exact string LBAND.
- The wildcarded string ’*BAND*’ will be used as a string pattern for matching. This will match any string which has the sub-string BAND in it.
- The string ’"*BAND*"' will also be used as a string pattern, matching any string which has the sub-string BAND in it.
- The string ’/.*BAND.+/' will be used as a regular expression. This will also match any string which has the sub-string BAND in it. (the .+ regex operator has the same meaning as the * wildcard operator of patterns).

2.3.2 The field Parameter

The `field` parameter is a string that specifies which field names or ids will be processed in the task or tool. The field selection expression consists of comma separated list of field specifications inside the string.

Field specifications can be literal field names, regular expressions or patterns (see § 2.3.1.1). Those fields for which the entry in the NAME column of the FIELD MS sub-table match the literal field name/regular expression/pattern are selected. If a field name/regular expression/pattern fails to match any field name, the given name/regular expression/pattern are matched against the field code. If still no field is selected, an exception is thrown.

Field specifications can also be given by their integer IDs. IDs can be a single or a range of IDs. Field ID selection can also be done as a boolean expression. For a field specification of the form '>ID', all field IDs greater than ID are selected. Similarly for '<ID' all field IDs less than the ID are selected.

For example, if the MS has the following observations:

MS summary:

```
=====
```

FIELDID	SPWID	NChan	Pol	NRows	Source Name
0	0	127	RR	10260	0530+135
1	0	127	RR	779139	05582+16320
2	0	127	RR	296190	05309+13319
3	0	127	RR	58266	0319+415
4	0	127	RR	32994	1331+305
5	1	1	RR,RL,LL,RR	23166	KTIP

one might select

```
field = '0~2,KTIP'      # FIELDID 0,1,2 and field name KTIP
field = '0530+135'     # field 0530+135
field = '05*'          # fields 0530+135,05582+16320,05309+13319
```

2.3.3 The spw Parameter

The `spw` parameter is a string that indicates the specific spectral windows and the channels within them to be used in subsequent processing. Spectral window selection ('SPWSEL') can be given as a spectral window integer ID, a list of integer IDs, a spectral window name specified as a literal string (for exact match) or a regular expression or pattern.

The specification can be via frequency ranges or by indexes. A range of frequencies are used to select all spectral windows which contain channels within the given range. Frequencies can be specified with an optional unit — the default unit being Hz. Other common choices for radio and mm/sub-mm data are kHz, MHz, and GHz. You will get the entire spectral windows, not just the channels in the specified range. You will need to do channel selection (see below) to do that.

The `spw` can also be selected via comparison for integer IDs. For example, '>ID' will select all spectral windows with ID greater than the specified value, while '<ID' will select those with ID lesser than the specified value.

Spectral window selection using strings follows the standard rules:

```
spw = '1'              # SPWID 1
spw = '1,3,5'         # SPWID 1,3,5
spw = '0~3'           # SPWID 0,1,2,3
spw = '0~3,5'         # SPWID 0,1,2,3 and 5
spw = '<3,5'           # SPWID 0,1,2,3 and 5
spw = '*'             # All spectral windows
spw = '1412~1415MHz' # Spectral windows containing 1412-1415MHz
```

In some cases, the spectral windows may allow specification by name. For example,

```
spw = '3mmUSB, 3mmLSB'      # choose by names (if available)
```

might be meaningful for the dataset in question.

Note that the order in which multiple `spw`s are given may be important for other parameters. For example, the `mode = 'channel'` in `clean` uses the first `spw` as the origin for the channelization of the resulting image cube.

2.3.3.1 Channel selection in the `spw` parameter

Channel selection can be included in the `spw` string in the form `'SPWSEL:CHANSEL'` where `CHANSEL` is the channel selector. In the end, the spectral selection within a given spectral window comes down to the selection of specific channels. We provide a number of shorthand selection options for this. These `CHANSEL` options include:

- *Channel ranges*: `'START~STOP'`
- *Frequency ranges*: `'FSTART~FSTOP'`
- *Channel striding/stepping*: `'START~STOP^STEP'` or `'FSTART~FSTOP^FSTEP'`

The most common selection is via channel ranges `'START~STOP'` or frequency ranges `'FSTART~FSTOP'`:

```
spw = '0:13~53'           # spw 0, channels 13-53, inclusive
spw = '0:1413~1414MHz'   # spw 0, 1413-1414MHz section only
```

All ranges are inclusive, with the channel given by, or containing the frequency given by, `START` and `STOP` plus all channels between included in the selection. You can also select the spectral window via frequency ranges `'FSTART~FSTOP'`, as described above:

```
spw = '1413~1414MHz:1413~1414MHz' # channels falling within 1413~1414MHz
spw = '*:1413~1414MHz'           # does the same thing
```

You can also specify multiple spectral window or channel ranges, e.g.

```
spw = '2:16, 3:32~34'         # spw 2, channel 16 plus spw 3 channels 32-34
spw = '2:1~3;57~63'          # spw 2, channels 1-3 and 57-63
spw = '1~3:10~20'            # spw 1-3, channels 10-20
spw = '*:4~56'               # all spw, channels 4-56
```

Note the use of the wildcard in the last example.

A step can also be included using `'^STEP'` as a postfix:

```
spw = '0:10~100^2'           # chans 10,12,14,...,100 of spw 0
spw = ':^4'                  # chans 0,4,8,... of all spw
spw = ':100~150GHz^10GHz'    # closest chans to 100,110,...,150GHz
```

A step in frequency will pick the channel in which that frequency falls, or the nearest channel.

2.3.4 The selectdata Parameters

The `selectdata` parameter, if set to `True` (default), will expand the inputs to include a number of sub-parameters, given below and in the individual task descriptions (if different). If `selectdata = False`, then the sub-parameters are treated as blank for selection by the task.

The common `selectdata` expanded sub-parameters are:

2.3.4.1 The antenna Parameter

The `antenna` selection string is a semi-colon (';') separated list of baseline specifications. A baseline specification is of the form:

- 'ANT1' — Select all baselines including the antenna(s) specified by the selector ANT1.
- 'ANT1&' — Select only baselines between the antennas specified by the selector ANT1.
- 'ANT1&ANT2' — Select only the cross-correlation baselines between the antennas specified by selector ANT1 and antennas specified by selector ANT2. Thus 'ANT1&' is an abbreviation for 'ANT1&ANT1'.
- 'ANT1&&ANT2' — Select only auto-correlation and cross-correlation baselines between antennas specified by the selectors ANT1 and ANT2. Note that this is what the default `antenna=''` gives you.
- 'ANT1&&&' — Select only autocorrelations specified by the selector ANT1.

The selectors ANT1 and ANT2 are comma-separated lists of antenna integer-IDs or literal antenna names, patterns, or regular expressions. The ANT strings are parsed and converted to a list of antenna integer-IDs or IDs of antennas whose name match the given names/pattern/regular expression. Baselines corresponding to all combinations of the elements in lists on either side of ampersand are selected.

Integer IDs can be specified as single values or a range of integers. When items of the list are parsed as literal strings or regular expressions or patterns (see § 2.3.1 for more details on strings). All antenna names that match the given string (exact match)/regular expression/pattern are selected.

ALERT: Just for antenna selection, a user supplied integer (or integer list) is converted to a string and matched against the antenna name. If that fails, the normal logic of using an integer as an integer and matching it with antenna index is done. Note that currently there is no method for specifying a pure index (e.g. a number that will not first be checked against the name).

The comma is used only as a separator for the list of antenna specifications. The list of baselines specifications is a semi-colon separated list, e.g.

```
antenna = '1~3 & 4~6 ; 10&11'
```

will select baselines between antennas 1,2,3 and 4,5,6 ('1&4', '1&5', ..., '3&6') plus baseline '10&11'.

The wildcard operator ('*') will be the most often used pattern. To make it easy to use, the wildcard (and only this operator) can be used without enclosing it in quotes. For example, the selection

```
antenna = 'VA*'
```

will match all antenna names which have 'VA' as the first 2 characters in the name (irrespective of what follows after these characters).

There is also a negation operator "!" that can be used to de-select antennas or baselines.

Some examples:

```
antenna=' '          # shows blank autocorr pages
antenna='*&*'       # does not show the autocorrs
antenna='*&&*'       # show both auto and cross-cor (default)
antenna='*&&&*'       # shows only autocorrs

antenna='5&*'        # shows non-auto baselines with AN 5

antenna='5,6&&&*'     # AN 5 and 6 autocor
antenna='5&&&&;6&*'    # AN 5 autocor plus cross-cors to AN 6

antenna='!5'         # baselines not involving AN 5
```

Antenna numbers as names: Needless to say, naming antennas such that the names can also be parsed as a valid token of the syntax is a bad idea. Nevertheless, antenna names that contain any of the reserved characters and/or can be parsed as integers or integer ranges can still be used by enclosing the antenna names in double quotes (' "ANT" '). E.g. the string

```
antenna = '10~15,21,VA22'
```

will expand into an antenna ID list 10,11,12,13,14,15,21,22 (assuming the index of the antenna named 'VA22' is 22). If, however, the antenna with ID index 50 is named '21', then the string

```
antenna = '10~15,21,VA22'
```

will expand into an antenna ID list of 10,11,12,13,14,15,50,22. *Keep in mind that numbers are FIRST matched against names, and only against indices if that matching fails.* There is currently no way to force a selection to use the index, and if there an antenna with that name it will select that.

Read elsewhere (e.g. `info regex` under Unix) for details of regular expression and patterns.

Antenna stations Instead of antenna names, the antenna station names are also accepted by the selection syntax., e.g. 'N15' for the JVLA.

ANT@STATION sections syntax Sometimes, data from multiple array configurations are stored in a single MS. But some antennas may have been moved during reconfiguration and the 'ANT@STATION' syntax can distinguish between them. 'ANT' is the antenna name or index and 'STATION' is the antenna station name, e.g., 'EA12@W03' selects antenna EA012 but only at times when it is positioned on station W03. Wildcards are accepted, e.g. 'EA12@*' selects all visibilities from antenna EA12, and '*@W03' would select all antennas that are located on station 'W03' during any observations included in the MS.

2.3.4.2 The scan Parameter

The `scan` parameter selects the scan ID numbers of the data. There is currently no naming convention for scans. The scan ID is filled into the MS depending on how the data was obtained, so use this with care.

Examples:

```
scan = '3'           # scan number 3.
scan = '1~8'        # scan numbers 1 through 8, inclusive
scan = '1,2,4,6'    # scans 1,2,4,6
scan = '<9'          # scans <9 (1-8)
```

NOTE: ALMA and VLA/JVLA number scans starting with 1 and not 0. You can see what the numbering is in your MS using the `listobs` task with `verbose=True` (see § 2.2.7).

2.3.4.3 The timerange Parameter

The time strings in the following (T0, T1 and dT) can be specified as YYYY/MM/DD/HH:MM:SS.FF. The time fields (i.e., YYYY, MM, DD, HH, MM, SS and FF), starting from left to right, may be omitted and they will be replaced by context sensitive defaults as explained below.

Some examples:

1. `timerange='T0~T1'`: Select all time stamps from T0 to T1. For example:

```
timerange = '2007/10/09/00:40:00 ~ 2007/10/09/03:30:00'
```

Note that fields missing in T0 are replaced by the fields in the time stamp of the first valid row in the MS. For example,

```
timerange = '09/00:40:00 ~ 09/03:30:00'
```

where the YY/MM/ part of the selection has been defaulted to the start of the MS.

Fields missing in T1, such as the date part of the string, are replaced by the corresponding fields of T0 (after its defaults are set). For example:

```
timerange = '2007/10/09/22:40:00 ~ 03:30:00'
```

does the same thing as above.

2. `timerange='T0'`: Select all time stamps that are within an integration time of T0. For example,

```
timerange = '2007/10/09/23:41:00'
```

Integration time is determined from the first valid row (more rigorously, an average integration time should be computed). Default settings for the missing fields of T0 are as in (1).

3. `timerange='T0+dT'`: Select all time stamps starting from T0 and ending with time stamp T0+dT. For example,

```
timerange = '23:41:00+01:00:00'
```

picks an hour-long chunk of time.

Defaults of T0 are set as usual. Defaults for dT are set from the time corresponding to MJD=0. Thus, dT is a specification of length of time from the assumed nominal "start of time".

4. `timerange='>T0'`: Select all times greater than T0. For example,

```
timerange = '>2007/10/09/23:41:00'
timerange = '>23:41:00'           # Same thing without day specification
```

Default settings for T0 are as above.

5. `timerange='<T1'`: Select all times less than T1. For example,

```
timerange = '<2007/10/09/23:41:00'
```

Default settings for T1 are as above.

An ultra-conservative selection might be:

```
timerange = '1960/01/01/00:00:00~2020/12/31/23:59:59'
```

which would choose all possible data!

2.3.4.4 The `uvrange` Parameter

Rows in the MS can also be selected based on the uv-distance or physical baseline length that the visibilities in each row correspond to. This `uvrange` can be specified in various formats.

The basic building block of uv-distance specification is a valid number with optional units in the format N[UNIT] (the unit in square brackets is optional). We refer to this basic building block as `UVDIST`. The default unit is meter. Units of length (such as 'm' and 'km') select physical

baseline distances (independent of wavelength). The other allowed units are in wavelengths (such as 'lambda', 'klambda' and 'Mlambda' and are true uv-plane radii

$$r_{uv} = \sqrt{u^2 + v^2}. \quad (2.1)$$

If only a single UVDIST is specified, all rows, the uv-distance of which exactly matches the given UVDIST, are selected.

UVDIST can be specified as a range in the format 'N0~N1[UNIT]' (where N0 and N1 are valid numbers). All rows corresponding to uv-distance between N0 and N1 (inclusive) when converted the specified units are selected.

UVDIST can also be selected via comparison operators. When specified in the format '>UVDIST', all visibilities with uv-distances greater than the given UVDIST are selected. Likewise, when specified in the format '<UVDIST', all rows with uv-distances less than the given UVDIST are selected.

Any number of above mentioned uv-distance specifications can be given as a comma-separated list.

Examples:

```
uvrange = '100~200km'           # an annulus in physical baseline length
uvrange = '24~35Mlambda, 40~45Mlambda' # two annuli in units of mega-wavelengths
uvrange = '< 45klambda'         # less than 45 kilolambda
uvrange = '> 0lambda'           # greater than zero length (no auto-corrs)
uvrange = '100km'              # baselines of length 100km
uvrange = '100klambda'         # uv-radius 100 kilolambda
```

2.3.4.5 The correlation Parameter

The `correlation` parameter will select between different correlation products. They can be either the correlation ID or values such as 'XX', 'YY', 'XY', 'YX', 'RR', 'LL', 'RL', 'LR'.

2.3.4.6 The intent Parameter

`intent` is the scan intent that was specified when the observations were set up. They typically describe what was intended with a specific scan, i.e. a flux or phase calibration, a bandpass, a pointing, an observation of your target, or something else or a combination. The format for the scan intents of your observations are listed in the logger when you run `listobs`. Minimum matching with wildcards will work, like '*BANDPASS*'. This is especially useful when multiple intents are attached to scans.

2.3.4.7 The observation Parameter

The `observation` parameter can select between different observation IDs. They will be assigned to parts of a combined data set during a run of `concat`. Each input MS will receive its own observation id in the process.

2.3.4.8 The `mselect` Parameter

More complicated selections within the MS structure are possible using the Table Query Language (TaQL). This is accessed through the `mselect` parameter.

Note that the TaQL syntax does not follow the rules given in § 2.3.1 for our other selection strings. TaQL is explained in more detail in **Aips++ NOTE 199 — Table Query Language** (<http://aips2.nrao.edu/docs/notes/199/199.html>). This will eventually become a CASA document. The specific columns of the MS are given in the most recent MS specification document: **Aips++ NOTE 229 — Measurement‘ Set definition version 2.0** (<http://aips2.nrao.edu/docs/notes/229/229.html>). This documentation will eventually be updated to the CASA document system.

Most selection can be carried out using the other selection parameters. However, these are merely shortcuts to the underlying TaQL selection. For example, field and spectral window selection can be done using `mselect` rather than through `field` or `spw`:

```
mselect='FIELD_ID == 0'           # Field id 0 only
mselect='FIELD_ID <= 1'         # Field id 0 and 1
mselect='FIELD_ID IN [1,2]'     # Field id 1 and 2
mselect='FIELD_ID==0 && DATA_DESC_ID==3' # Field id 0 in spw id 3 only
```

ALERT: The `mselect` style parameters will be phased out of the tasks. TaQL selection will still be available in the Toolkit.

Chapter 3

Data Examination and Editing

3.1 Plotting and Flagging Visibility Data in CASA

The tasks available for plotting and flagging of data are:

- `flagmanager` — manage versions of data flags (§ 3.2)
- `plotms` — create X-Y plots of data in MS and calibration tables, flag data (§ 3.3.1)
- `plotxy` — older X-Y plotter with some functionalities not yet implemented in `plotms` (§ 3.3.2)
- `flagdata` --- Data Flagging (§ 3.4)
- `flagcmd` --- manipulate and apply flags using `FLAG_CMD` table (§ 3.5)
- `browsetable` --- browse data in any CASA table (including a MS) (§ 3.6)
- `plotants` --- create simple plots of antenna positions (§ 3.3.3)
- `plotuv` --- plotting of uv-coverages (§ 3.3.4)

The following sections describe the use of these tasks.

Information on other related operations can be found in:

- `listobs` — list summary of a MS (§ 2.2.7)
- `listvis` — list data in a MS (§ 2.2.9)
- `selectdata` — general data selection syntax (§ 2.3)
- `viewer` — use the `casaviewer` to display the MS as a raster image, and flag it (§ 7)

3.2 Managing flag versions with flagmanager

The `flagmanager` task will allow you to manage different versions of flags in your data. These are stored inside a CASA flagversions table, under the name of the MS `<msname>.flagversions`. For example, for the MS `jupiter6cm.usecase.ms`, there will need to be `jupiter6cm.usecase.ms.flagversions` on disk. This is created on import (by `importvla` or `importuvfits`) or when flagging is first done on an MS without a `.flagversions` (e.g. with `plotxy`).

By default, when the `.flagversions` is created, this directory will contain a `flags.Original` in it containing a copy of the original flags in the `MAIN` table of the MS so that you have a backup. It will also contain a file called `FLAG_VERSION_LIST` that has the information on the various flag versions there. The `flagversions` are cumulative, i.e. a specific version number contains all the flags from the lower version numbers, too.

The inputs for `flagmanager` are:

```
vis          =          ''          # Name of input visibility file (MS)
mode        =          'list'       # Flag management operation (list,save,restore,delete)
```

The `mode='list'` option will list the available flagversions from the `jmsname.j.flagversions` file. For example:

```
CASA <102>: default('flagmanager')
CASA <103>: vis = 'jupiter6cm.usecase.ms'
CASA <104>: mode = 'list'
CASA <105>: flagmanager()
MS : /home/imager-b/smyers/Oct07/jupiter6cm.usecase.ms
```

```
main : working copy in main table
Original : Original flags at import into CASA
flagautocorr : flagged autocorr
xyflags : Plotxy flags
```

The `mode` parameter expands the options. For example, if you wish to save the current flagging state of `vis=jmsname.j`,

```
mode          =          'save'      # Flag management operation (list,save,restore,delete)
  versionname =          ''          # Name of flag version (no spaces)
  comment     =          ''          # Short description of flag version
  merge       =          'replace'   # Merge option (replace, and, or)
```

with the output version name specified by `versionname`. For example, the above `xyflags` version was written using:

```
default('flagmanager')
vis = 'jupiter6cm.usecase.ms'
mode = 'save'
versionname = 'xyflags'
comment = 'Plotxy flags'
flagmanager()
```

and you can see that there is now a sub-table in the flagversions directory

```
CASA <106>: ls jupiter6cm.usecase.ms.flagversions/
IPython system call: ls -F jupiter6cm.usecase.ms.flagversions/
flags.flagautocorr flags.Original flags.xyflags FLAG_VERSION_LIST
```

It is recommended that you use this facility regularly to save versions during flagging.

Note that if a flagversion already exists under a name, the task will give a warning and add a suffix `'old.timestamp'` to the previous version.

You can restore a previously saved set of flags using the `mode='restore'` option:

```
mode           = 'restore'      # Flag management operation (list,save,restore,delete)
  versionname  = ''             # Name of flag version (no spaces)
  merge        = 'replace'      # Merge option (replace, and, or)
```

The `merge` sub-parameter will control how the flags are restored. For `merge='replace'`, the flags in `versionname` will replace those in the MAIN table of the MS. For `merge='and'`, only data that is flagged in BOTH the current MAIN table and in `versionname` will be flagged. For `merge='or'`, data flagged in EITHER the MAIN or in `versionname` will be flagged.

The `mode='delete'` option can be used to remove `versionname` from the flagversions:

```
mode           = 'delete'      # Flag management operation (list,save,restore,delete)
  versionname  = ''             # Name of flag version (no spaces)
```

3.3 X-Y Plotting and Editing of the Data

There are three main X-Y plotting tasks in CASA:

- `plotms` — create X-Y plots of data in MS, flag data (§ 3.3.1)
- `plotxy` — older X-Y plotter with some functionalities not yet implemented in `plotms` (§ 3.3.2)
- `plotants` — create simple plots of antenna positions (§ 3.3.3)

3.3.1 MS Plotting and Editing using `plotms`

The principal way to get X-Y plots of visibility data and calibration tables is the `plotms` task. This task also provides editing capability. `Plotms` is a GUI-style plotter, based on Qt. It can either be started as a task within CASA (`plotms`) or from outside CASA (type `casaplotms` on the command line).

The current inputs to the `plotms` task are:

```

# plotms :: A plotter/interactive flagger for visibility data.
vis          =      ''      # input MS (or CalTable) (blank for none)
gridrows    =      1      # Number of subplot rows (default 1).
gridcols    =      1      # Number of subplot columns (default 1).
rowindex    =      0      # Row location of the plot (0-based, default 0)
colindex    =      0      # Column location of the plot (0-based, default 0)
plotindex   =      0      # Index to address a subplot (0-based, default 0)
xaxis       =      ''      # plot x-axis (blank for default/current)
yaxis       =      ''      # plot y-axis (blank for default/current)
selectdata  =      True    # data selection parameters
  field     =      ''      # field names or field index numbers (blank for all)
  spw       =      ''      # spectral windows:channels (blank for all)
  timerange =      ''      # time range (blank for all)
  uvrange   =      ''      # uv range (blank for all)
  antenna   =      ''      # antenna/baselines (blank for all)
  scan      =      ''      # scan numbers (blank for all)
  correlation =      ''    # correlations (blank for all)
  array     =      ''      # (sub)array numbers (blank for all)
  observation =      ''    # Select by observation ID(s)
  intent    =      ''      # Select observing intent
  msselect  =      ''      # MS selection (blank for all)

averagedata =      True    # data averaging parameters
  avgchannel =      ''      # average over channel? (blank = False, otherwise value
  # in channels)
  avgtime    =      ''      # average over time? (blank = False, other value in
  # seconds)
  avgscan    =      False   # only valid if time averaging is turned on. average
  # over scans?
  avgfield   =      False   # only valid if time averaging is turned on. average
  # over fields?
  avgbaseline =      False  # average over all baselines? (mutually exclusive with
  # avgantenna)
  avgantenna =      False  # average by per-antenna? (mutually exclusive with
  # avgbaseline)
  avgspw     =      False   # average over all spectral windows?
  scalar     =      False   # Do scalar averaging?

transform    =      False   # transform data in various ways?
extendflag  =      False   # have flagging extend to other data points?
iteraxis    =      ''      # the axis over which to iterate
customsymbol =      False  # set a custom symbol(s) for unflagged points
coloraxis   =      ''      # selects which data to use for colorizing
customflaggedsymbol =      False # set a custom plot symbol for flagged points
plotrange   =      []      # plot axes ranges: [xmin,xmax,ymin,ymax]
title       =      ''      # Title written along top of plot
xlabel      =      ''      # Text for horizontal axis. Blank for default.
ylabel      =      ''      # Text for vertical axis. Blank for default.
showmajorgrid =      False # Show major grid lines (horiz and vert.)
showminorgrid =      False # Show minor grid lines (horiz and vert.)
showlegend  =      False   # Show a legend on the plot.

```

```

plotfile      =      ''      # Name of plot file to save automatically.
showgui       =      True     # Show GUI
clearplots    =      True     # Remove any existing plots so new ones can replace
                                # them.

callib        =      ''      # Calibration library filename or dictionary for on-the-
                                # fly calibration.

```

Almost all of these parameters can also be set or modified from inside the `plotms` window. Note that, if the `vis` parameter is set to the name of a measurement or calibration table set here, when you start up `plotms`, the *entire* measurement set will be plotted, which can be time consuming. It is probably best to leave all parameters blank for now, setting them as needed inside the `plotms` GUI.

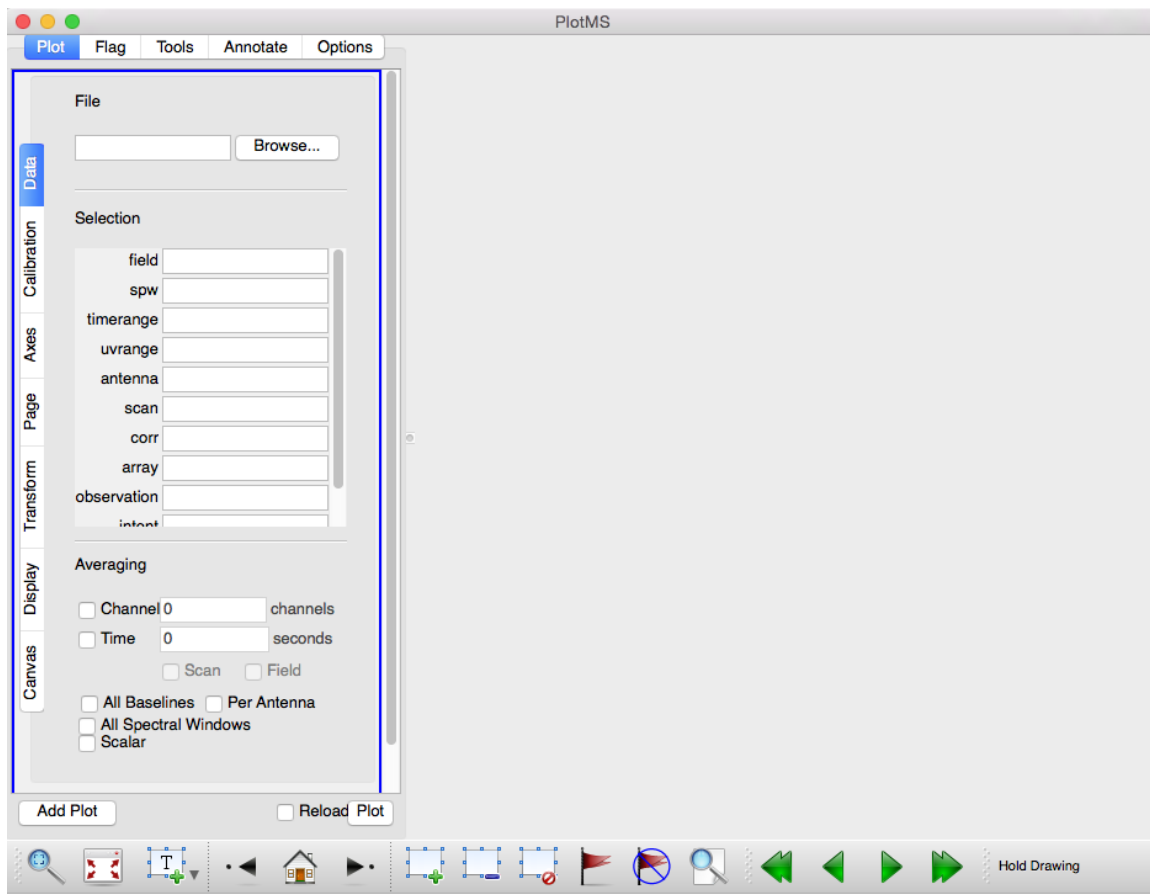


Figure 3.1: A freshly-started `plotms` GUI window. Note that the **Plots** > **Data** tab is selected, which is discussed in § 3.3.1.1, 3.3.1.7, and 3.3.1.9.

3.3.1.1 Loading and Selecting Data

When `plotms` is first started, a window will appear as in Figure 3.1. It will, by default, display the **Plots** tab (as chosen from the tabs at the top of the `plotms` window—e.g., **Plots, Flag, Tools...**) and the **Plots > Data** tab (as chosen from the tabs on the far left side of the `plotms` window—e.g., **Data, Calibration, Axes, Pages, Transform...**). First, a measurement set should be loaded by clicking on **Browse** near the top of the **Plots > Data** tab, and selecting a `.ms` directory (just select the directory itself; do not descend into the `.ms` directory). A plot can now be made of the measurement set by clicking on **Plot**—but beware, this would plot the entire measurement set, and could take quite some time! It is probably better to select a subset of the measurement set using the **Selection** windows in the **Plots > Data** tab before clicking **Plot**.

The options for data selection are:

- **field**
- **spw**
- **timerange**
- **uvrange**
- **antenna**
- **scan**
- **corr**
- **array**
- **observation**
- **msselect**

These are described in § 2.3. Note that, unlike when setting data selection parameters from the CASA command line, no quotation marks are needed around strings.

Once you have selected the desired subset of data, if you click **Plot**, `plotms` will by default plot amplitude versus time. See the next section for information about other possible axes.

For a given data selection, `plotms` will only load the data once. This speeds up plotting considerably when changing plot parameters such as different axes, colors etc. Sometimes, however, the data changes on disk, e.g., when other data processing tasks were applied. To force `plotms` to reload the data, checkmark the little **force reload** box left to the **Plot** button or press the *SHIFT* key while clicking the **Plot** button.

3.3.1.2 A Brief Note Regarding `plotms` Memory Usage

In order to provide a wide range of flexible interactive plotting options while minimizing the I/O burden, `plotms` caches the data values for the plot (along with a subset of relevant meta-info) in as efficient a manner as possible. For plots of large numbers of points, the total memory requirement can be quite large. `plotms` attempts to predict the memory it will require (typically 5 or 6 bytes per plotted point when only one axis is a data axis, depending upon the data shapes involved), and will complain if it believes there is insufficient memory to support the requested plot. For most practical interactive purposes (plots that load and draw in less than a few or a few 10s of minutes), there is usually not a problem on typical modern workstations (attempts to plot large datasets on small laptops might be more likely to encounter problems here).

The absolute upper limit on the number of simultaneously plotted points is currently set by the ability to index the points in the cache. For modern 64 bit machines, this is about 4.29 billion points (requiring around 25GB of memory). (Such plots are not especially useful interactively, since the I/O and draw become prohibitive.)

In general, it is usually most efficient to plot data in modest chunks of not more than a few hundred million points or less, either using selection or averaging. Note that all iterations are (currently) cached simultaneously for iterated plots, so iteration is not a way to manage memory use. A few hundred million points tends to be the practical limit of interactive `plotms` use w.r.t. information content and utility in the resulting plots, especially when you consider the number of available pixels on your screen.

In scripts, or for very large data sets, it can be desirable to use `plotms` in a non-interactive mode. This can be done by setting `showgui=False` and to directly plot into a png image specified by `plotfile`.

3.3.1.3 Plot Axes

The X and Y axes of a plot are selected by clicking on the **Plots** > **Axes** tab on the left side of the `plotms` window, and choosing an entry from the drop-down menus below **X Axis** and **Y Axis** (see Figure 3.2). Possible axes are:

- **Scan** — The scan number, as listed by `listobs` (§ 2.2.7) or the data summary in `plotms` (§ 3.3.1.9).
- **Field** — The field number, as listed by `listobs` (§ 2.2.7) or the `plotms` data summary (§ 3.3.1.9).
- **Time** — The time at which the visibility was observed, given in terms of calendar year (yyyy/mm/dd/hh:mm:ss.ss).
- **Interval** — The integration time in seconds.
- **Spw** — The spectral window number. The characteristics of each spectral window are listed in `listobs` (§ 2.2.7) or the `plotms` data summary (§ 3.3.1.9).

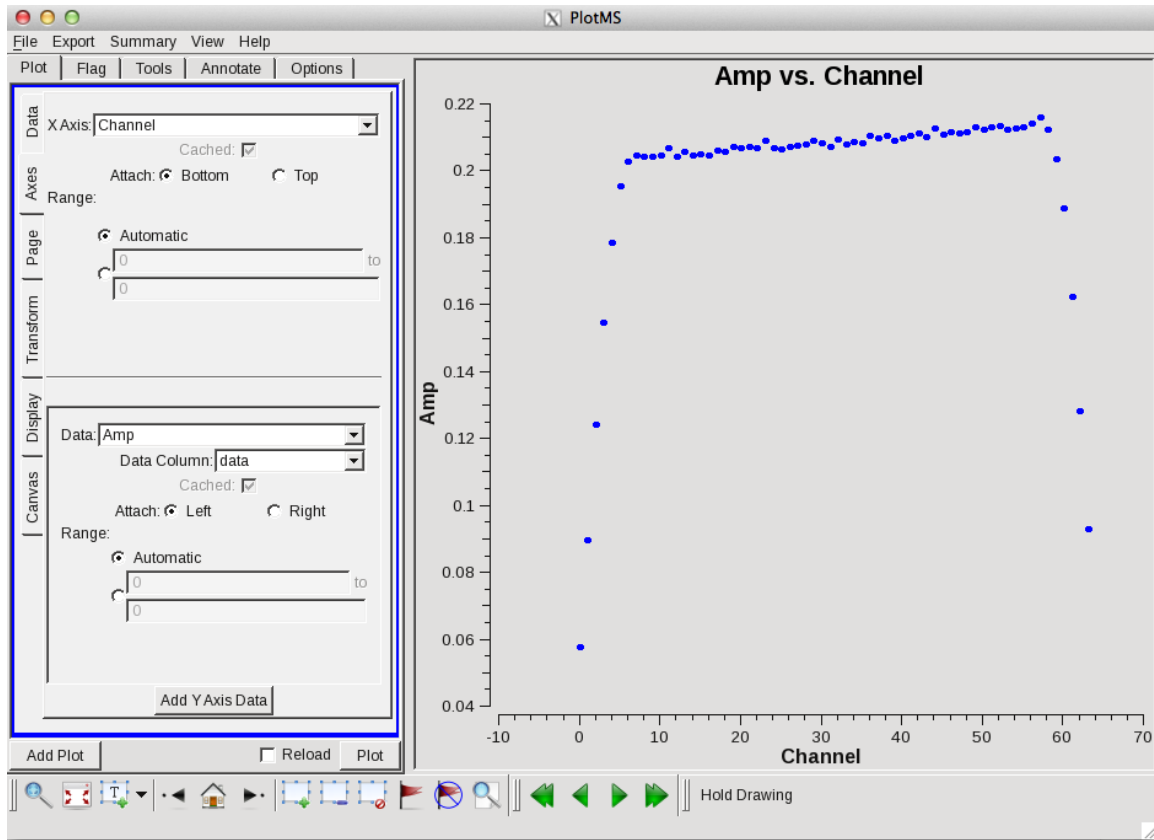


Figure 3.2: The **Plots > Axes** tab in the plotms GUI window, used to make a plot of **Amp** versus **Channel**.

- **Channel** — The spectral channel number.
- **Frequency** — Frequency in units of GHz. The frame for the frequency (e.g., topocentric, barycentric, LSRK) can be set in the **Plots > Trans** tab (§ 3.3.1.10).
- **Velocity** — Velocity in units of km s^{-1} , as defined by the **Frame**, **Velocity Defn**, and **Rest Freq** parameters in the **Plots > Transform** tab (§ 3.3.1.10).
- **Corr** — Correlations which have been assigned integer IDs: 5 = RR; 6 = RL; 7 = LR; and 8 = LL.
- **Antenna1** — The first antenna in a baseline pair; for example, for baseline 2-4, Antenna1 = 2. Antennae are numbered according to the antenna IDs listed in `listobs` (§ 2.2.7) or the plotms data summary (§ 3.3.1.9).
- **Antenna2** — The second antenna in a baseline pair; for baseline 2-4, Antenna2 = 4. Antennae are numbered according to the antenna IDs listed in `listobs` (§ 2.2.7) or the plotms data summary (§ 3.3.1.9).

- **Antenna** — Antenna ID for plotting antenna-based quantities. Antennae are numbered according to the antenna IDs listed in `listobs` (§ 2.2.7) or the `plotms` data summary (§ 3.3.1.9).
- **Baseline** — The baseline number.
- **Row** — The MS data row number.
- **Real** and **Imag** — The real and imaginary parts of the visibility in units which are proportional to Jansky (for data which are fully calibrated, the units should be Jy).
- **UVDist** — Projected baseline separations in units of meters. Note that **UVDist** is *not* a function of frequency.
- **UVwave** — Projected baseline separations in units of the observing wavelength (λ , not kilolambda). **UVDist_L** is a function of frequency, and therefore, there will be a different data point for each frequency channel.
- **U**, **V**, and **W** — u , v , and w in units of meters.
- **Uwave**, **Vwave**, and **Wwave** — u , v , and w in units of wavelengths λ .
- **Amp** — Data amplitudes in units which are proportional to Jansky (for data which are fully calibrated, the units should be in Jy).
- **Phase** — Data phases in units of degrees.
- **Wt** and **Wt*Amp** — the weight of the visibility (see Appendix F) and the product of the weight and the amplitude
- **WtSp** — WEIGHT_SPECTRUM column, i.e. a weight per channel
- **Sigma** — the SIGMA column of the visibilities (see Appendix F)
- **SigmaSp** — SIGMA_SPECTRUM column, i.e. a SIGMA per channel
- **Flag** and **FlagRow** — Data which are flagged have `Flag = 1`, whereas unflagged data are set to `Flag = 0`, **FlagRow** is the MS row number. Note that, to display flagged data, you will have to click on the **Plots > Display** tab and choose a **Flagged Points Symbol** (§ 3.3.1.8).
- **Azimuth** and **Ant-Azimuth** — Azimuth in units of degrees. **Azimuth** plots a fiducial value for the entire array, while **Ant-Azimuth** plots the azimuth for each individual antenna (their azimuths will differ by small amounts, because each antenna is located at a slightly different longitude, latitude, and elevation).
- **Elevation** and **Ant-Elevation** — Elevation in units of degrees. **Elevation** is a representative value for the entire array, while **Ant-Elevation** is the elevation for each individual antenna (their elevations will differ by small amounts, because each antenna is located at a slightly different longitude, latitude, and elevation).
- **HourAngle** — Hour angle in units of hours. This is a fiducial value for the entire array.

- **ParAngle** and **Ant-ParAng** — Parallaxic angle in units of degrees. **ParAngle** is the fiducial parallaxic angle for all antennae in the array, while **Ant-ParAng** plots the parallaxic angle for each individual antenna (their parallaxic angles will differ by small amounts, because each antenna is located at a slightly different longitude, latitude, and elevation).
- **Row** — Data row number. A row number corresponds to a unique time, baseline, and spectral window in the measurement set.
- **FlagRow** — In some tasks, if a whole data row is flagged, then FlagRow will be set to 1 for that row. Unflagged rows have FlagRow = 0. However, note that some tasks (like `plotms`) may flag a row, but *not* set FlagRow = 1. It is probably better to plot Flag than FlagRow for most applications.
- **GainAmp**, **GainPhase**, **GainReal**, **GainImag** — are the amplitude, phase, real and imaginary part of the calibration tables for regular complex gain tables.
- **Delay** — The delay of a delay calibration table
- **Opac** — Opacity values of a Opacity calibration table
- **SwPower** — Switched Power values for VLA switched power calibration tables
- **Tsys** — Tsys for ALMA Tsys calibration tables

If the data axis selected from the drop-down menu is already stored in the cache (therefore implying that plotting will proceed relatively quickly), an “X” will appear in the checkbox next to **In Cache?**. If the data shall be reloaded from disk, the “force reload” checkmark should be set at the bottom of this display.

For relevant data axes like **Amp** and **Phase**, the user will be presented with the option to plot raw data or calibrated data. This can be selected via a drop-down menu called **Data Column**, located directly under the drop-down menu for X or Y Axis selection (see the Y axis in Figure 3.2). To plot raw data, select “data”; to plot calibrated data, select “corrected”. Note that this choice will only have an impact on a plot if a calibration table has been applied to the measurement set (see `applycal`, Sect. 4.6.1).

If a data model has been applied to the measurement set (e.g., with `setjy`, Sect. 4.3.5) it can be plotted by selecting “model” from the **Data Column** menu. Residuals can be plotted via “corrected-model”, “data-model”, “data/model”, and “corrected/model”.

3.3.1.4 Calibration Library

The keyword `'callib'` in the `plotms` task parameters can be used to provide a calibration library file (see Appendix G). This file allows one to apply calibration tables to the uncalibrated data on the fly, i.e. without a run of `applycal` beforehand. The tab *Calibration* on the right hand side contains a field to specify the calibration library or to specify the calibration library commands directly. There’s also a switch to either apply the calibration library to produce the “corrected” data (“Calibration On”) or to show an existing “corrected” data column (“Calibration Off”).

3.3.1.5 Tools

Various tools—selectable as icon buttons at the bottom of the `plotms` window—can be used to zoom, edit, and locate data. The icon buttons can be seen at the bottom of Figures 3.1 and 3.2, and are, from left to right:

- **Zoom** — The “magnifying glass” button (1st on left) lets you draw a box around a region of the plot (left-click on one corner of the box, and drag the mouse to the opposite corner of the desired box), and then zooms in on this box.
- **Pan** — The “four-arrow” button (2nd from left) lets you pan around a zoomed plot.
- **Annotate** — The 3rd button from the left is chosen from a drop-down menu to either **Annotate Text** (“T with a green diamond” button) or **Annotate Rectangle** (“pencil” button). In the **Annotate Text** environment, click on a location in the plot where text is desired; a window will pop up, allowing you to type text in it. When you click the **OK** button, this text will appear on the plot. **Annotate Rectangle** simply lets you draw a box on the plot by left-clicking and dragging the mouse. By clicking on the **Annotator** tab near the top of the `plotms` window, different fonts, colors, line styles, etc. can be selected for annotations.
- **Home** — The “house” button (5th from left) returns to the original zoom level.
- **Stack Back** and **Stack Forward** — The left and right arrow buttons (4th and 6th from left) step through the zoom settings you’ve visited.
- **Mark Regions** — The “box with a green diamond” button (7th from left) lets you mark a region for flagging, unflagging, or locating. Left-click on one corner of the desired region, and then drag the mouse to set the opposite corner of the region. You can mark multiple boxes before performing an operation on them.
- **Clear Regions** — Clicking on the “box with a red circle” button (8th from left) will clear *all* regions which have been marked using **Mark Regions**.
- **Flag** — Click on the “flag” button (9th from left) to flag all points in the marked regions.
- **Unflag** — Click on the “crossed-out flag” button (10th from left) to unflag any flagged points that would be in the marked regions (even if invisible).
- **Locate** — The “magnifying glass on a sheet of paper” button (11th from left) will print out information to the command line about points in the marked regions.
- **Hold Drawing** — If the “hold drawing” button (rightmost, or 12th from left) is depressed, and if new plot axes are selected from the **Plots > Axes** tab, these new data will be cached but not plotted. When the button is clicked on again and un-depressed, it will automatically plot the data that was last requested. This can be particularly useful when changing the size of the `plotms` window.

Under the **Options** tab near the top of the `plotms` window one can select the layout of the page. For multiple plots per page, one can select the grid layout, ie the number of rows and columns that determine the number of sub-plots. The **when changing plot axes, clear any existing regions or annotations** checkbox determines when regions and annotation are deleted from the plot. The **Tool Button Style** drop-drop down menu determines if icons and/or text represent the buttons at the bottom of the `plotms` window.

It is possible to hide these icons by going to the **View > Toolbars** menu at the top of the `plotms` window and un-depressing the **Tools** option (except for **Hide Drawing**, which is hidden by clicking on **View > Toolbars > Display**). In addition, the above tools can also be accessed by clicking on the **Tools** tab near the top of the `plotms` window (just below the **View** menu).

The **Tools** tab also enables one additional tool, the **Tracker**. To use **Tracker**, click on the **Hover** and/or

Display checkbox, and place your mouse over the plot. **Tracker** will output the X and Y position of your mouse, either as text superimposed on the plot near your mouse (if **Hover** is selected) or in the blank window in the **Tools** tab (if **Display** is selected). Pressing the *SPACE* bar will copy the lines into the larger white box below to the right. This can be repeated many times and a log of positions and values will be created. The content in the box can then be easily copied and pasted into any other application that is used for data analysis. The **Clear** button wipes out the content of the box for a fresh start into new scientific adventures.

3.3.1.6 Interactive Flagging in `plotms`

Interactive flagging, on the principle of “see it — flag it”, is possible on the X-Y display of the data plotted by `plotms`. The user can use the cursor to mark one or more regions, and then flag, unflag, or list the data that falls in these zones of the display.

Using the row of icons buttons at the bottom of the `plotms` window (§ 3.3.1.5), click on the **Mark Regions** button (which will appear to depress), then mark a region by left-clicking and dragging the mouse (each click and drag will mark an additional region). You can get rid of all your regions by clicking on the **Clear Regions**. Once regions are marked, you can then click on one of the other buttons to take action:

1. **Flag** — flag the points in the region(s),
2. **Unflag** — unflag flagged points in the region(s),
3. **Locate** — spew out a list of the points in the region(s) to the command line (Warning: this could be a long list!).

Figure 3.3 shows an example of marking regions and then clicking the **Flag** button. Whenever you click on a button, that action occurs without requiring an explicit disk-write. If you quit `plotms` and re-enter, you will see your previous edits.

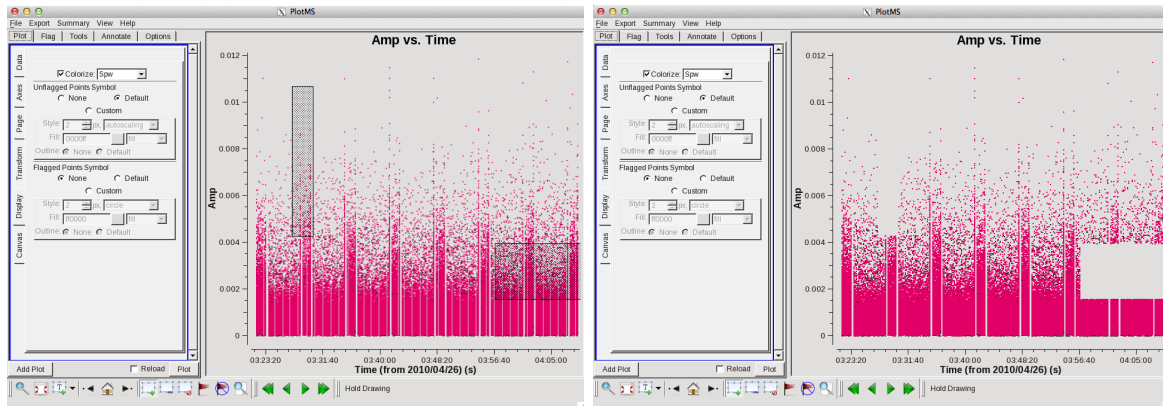


Figure 3.3: Plot of amplitude versus time, before (left) and after (right) flagging two marked regions. To unflag these regions, mark the two same regions and click the **Unflag** button.

A table with the name $jmsname.j.flagversions$ (where $vis=<msname>$) will be created in the same directory if it does not exist already. It is recommended that you save important flagging stages using the **flagmanager** task (§ 3.2).

Flags can also be extended with options in the **Flagging** tab, found near the top of the **plotms** window. Flag extension enables the user to plot a subset of the data and extend the flagging to a wider set. In this release, the only functional extensions are over channel and correlation.

By checking the boxes next to **Extend Flags** and **Channel**, flagging will be extended to other channels in the same *spw* as the displayed point. For example, if *spw*='0:0' and channel 0 is displayed, then flagging will extend to all channels in *spw* 0.

By checking the boxes next to **Extend Flags** and **Correlation**, flags will be extended beyond the correlations displayed. Currently the only option is to extend to **All** correlations, implying that all correlations will be flagged, e.g. with RR displayed, the correlations RR, RL, LR, and LL will all be flagged.

WARNING: use of flag extensions may lead to deletion of much more data than desired. Be careful!

3.3.1.7 Averaging Data

The **Plots > Data** tab enables averaging of the data in order to increase signal-to-noise of the plotted points or to increase plotting speed. The options for **Averaging** are:

- **channel**
- **time**
- **all baselines or per antenna**

- all spectral windows
- scalar

The box next to a given **Averaging** mode needs to be checked for that averaging to take effect.

For example, to average n channels together, the user would click on the box next to **Channels** so that an “X” appears in it, and then type the number n in the empty box. When the user next clicks on **Plot**, every n channels will then be averaged together and the total number of channels plotted will be decreased by a factor of n .

Time averaging is a little trickier, as it is controlled by three fields. If the checkbox next to **Time** under **Averaging** is clicked on, a blank box with units of seconds will become active, along with two additional checkboxes: **Scan** and **Field**. If averaging is desired over a relatively short interval (say, 30 seconds, shorter than the scan length), a number can simply be entered into the blank box and, when the data are replotted, the data will be time averaged. Clicking on the **Scan** or **Field** checkbox in this case will have no impact on the time averaging.

These checkboxes become relevant if averaging over a relatively long time—say the entire observation, which consists of multiple scans—is desired. Regardless of how large a number is typed into the **Time** averaging blank box, only data within individual scans will be averaged together. In order to average data across scan boundaries, the **Scan** checkbox must be clicked on and the data replotted. Finally, clicking on the **Field** checkbox enables the averaging of multiple fields together in time.

Clicking on the **All Baselines** checkbox will average all baselines in the array together. Alternatively, the **Per Antenna** box may be checked, which will average all baselines for a given antenna together. In this case, all baselines are represented twice; baseline 3-24 will contribute to the averages for both antenna 3 and antenna 24. This can produce some rather strange-looking plots if the user also selects on antenna—say, if the user requests to plot only antenna 0 and then averages **Per Antenna**. In this case, an average of all baselines including antenna 0 will be plotted, but each individual baseline including antenna 0 will also be plotted (because the presence of baselines 0-1, 0-2, 0-3, etc. trigger **Per Antenna** averaging to try and compute averages for antennae 1, 2, 3, etc. Therefore, baseline 0-1 will contribute to the average for antenna 0, but it will also singlehandedly be the average for antenna 1.)

Spectral windows can be averaged together by checking the box next to **All Spectral Windows**. This will result in, for a given channel n , all channels n from the individual spectral windows being averaged together.

Finally, the default mode is vector averaging, where the complex average is formed by averaging the real and imaginary parts of the relevant visibilities. If **Scalar** is chosen, then the amplitude of the average is formed by a scalar average of the individual visibility amplitudes.

When averaging, `plotms` will prefer unflagged data. I.e., if an averaging bin contains any unflagged data at all, only the average of the unflagged will be shown. For averaging bins that contain *only* unflagged data, the average of that unflagged data will be shown. When flagging on a plot of averaged data, the flags will be applied to the unaveraged data in the MS.

3.3.1.8 Plot Symbols

Plot symbols are selected in the **Plots > Display** tab. Most fundamentally, the user can choose to plot unflagged data and/or flagged data. By default, unflagged data is plotted (the circle next to **Default** is checked under **Unflagged Points Symbol**), and flagged data is not plotted (the circle next to **None** is checked under **Flagged Points Symbol**). We note here that plotting flagged data on an averaged plot is undertaken at the user's own risk, as the distinction between flagged points and unflagged points becomes blurred if data are averaged over a dimension that is partially flagged. Take, for example, a plot of amplitude versus time where all channels are averaged together, but some channels have been flagged due to RFI spikes. In creating the average, **plotms** will skip over the flagged channels and only use the unflagged ones. The averaged points will be considered unflagged, and the flagged data will not appear on the plot at all.

A selection of **None** produces no data points, **Default** results in data points which are small circles (blue for unflagged data and red for flagged data), and **Custom** allows the user to define a plot symbol. If **Custom** plot symbols are chosen, the user can determine the symbol size by typing a number in the blank box next to **px** or by clicking on the adjacent up or down arrows. Symbol shape can be chosen from the drop-down menu to be either "circle", "square", "diamond", or "pixel" (note that "pixel" only has one possible size). "autoscaling" attempts to adjust the size of the points from dots to circles of different sizes, depending on how many points are plotted. Symbol color can be chosen by typing a hex color code in the blank box next to **Fill:** (e.g., "ff00ff"), or by clicking on the ... button and selecting a color from the pop-up GUI. The adjacent drop-down menu provides options for how heavily the plot symbol is shaded with this color, from heaviest to lightest: "fill", "mesh1", "mesh2", "mesh3", and "no fill". Finally, the plot symbol can be outlined in black (if **Outline: Default** is checked) or not (if **Outline: None** is checked). Note that if "no fill" and **Outline: None** are selected, the plot symbols will be invisible.

Finally, unflagged data points can be given informative symbol colors using the **Colorize** parameter. By checking the box next to **Colorize** and selecting a data dimension from the drop-down menu, the data will be plotted with colors that vary along that dimension. For example, if "corr" is chosen from the **Colorize** menu, "RR", "LL", "RL", and "LR" data will each be plotted with a different color. Note that, currently, **colorize** and plotting flagged data appear to be incompatible; a plot can only include one of these special features at a time.

3.3.1.9 Summarizing Data

Information about the measurement set can be obtained from within **plotms** by clicking on the **Summary** button, found at the top menu bar. If "All" is chosen from the pull-down menu next to **Summary**, **listobs**-style output about scans, correlator configurations, and antennae will be written to the command line from which **plotms** was started. For more detail, click on the **Verbose** checkbox. For a specific subset of the data, choose a selection from the pull-down menu like "Antenna" or "Field".

3.3.1.10 Defining Frequency and Velocity

If the user plans to plot **Frequency**, the reference frame must be defined. By default, the plotted frequency is simply that observed at the telescope. However, transformations can be made by choosing a **Frame** from the drop-down menu in the **Plots > Transform** tab. Frequency reference frames can be chosen to be:

- **LSRK** — local standard of rest (kinematic)
- **LSRD** — local standard of rest (dynamic)
- **BARY** — barycentric
- **GEO** — geocentric
- **TOPO** — topocentric
- **GALACTO** — galactocentric
- **LGROUP** — Local Group
- **CMB** — cosmic microwave background dipole

Velocity is affected by the user's choice of **Frame**, but it is also impacted by the choice of velocity definition and spectral line rest frequency. The velocity definition is chosen from the **Velocity Defn** drop-down menu in the **Plots > Trans** tab, offering selections of **Radio**, **True**, or **Optical**.

For more information on frequency frames and spectral coordinate systems, see the paper by Greisen et al. (A&A, 446, 747, 2006) ¹.

Finally, the spectral line's rest frequency in units of MHz should be typed into the blank box next to **Rest Freq** in the **Plots > Trans** tab. You can use the *me.spectralline* tool method to turn transition names into frequencies

```
CASA <16>: me.spectralline('HI')
Out[17]:
{'m0': {'unit': 'Hz', 'value': 1420405751.786},
 'refer': 'REST',
 'type': 'frequency'}
```

For a list of known lines in the CASA *measures* system, use the toolkit command *me.linelist()*. For example:

```
CASA <21>: me.linelist()
Out[21]: 'HI H186A H185A H184A H183A H182A H181A H180A H179A H178A H177A H176A H175A
H174A H173A H172A H171A H170A H169A H168A H167A H166A H165A H164A H163A H162A H161A H160A...
He182A He181A He180A He179A He178A He177A He176A He175A He174A He173A He172A He171A He170A
He169A He168A He167A He166A He165A He164A He163A He162A He161A He160A He159A He158A He157A...
```

¹Also at <http://www.aoc.nrao.edu/~egreisen/scs.ps>

```

C186A C185A C184A C183A C182A C181A C180A C179A C178A C177A C176A C175A C174A C173A C172A
C171A C170A C169A C168A C167A C166A C165A C164A C163A C162A C161A C160A C159A C158A C157A...
NH3_11 NH3_22 NH3_33 NH3_44 NH3_55 NH3_66 NH3_77 NH3_88 NH3_99 NH3_1010 NH3_1111 NH3_1212
OH1612 OH1665 OH1667 OH1720 OH4660 OH4750 OH4765 OH5523 OH6016 OH6030 OH6035 OH6049 OH13433
OH13434 OH13441 OH13442 OH23817 OH23826 CH3OH6.7 CH3OH44 H2O22 H2CO4.8 CO_1_0 CO_2_1 CO_3_2
CO_4_3 CO_5_4 CO_6_5 CO_7_6 CO_8_7 13CO_1_0 13CO_2_1 13CO_3_2 13CO_4_3 13CO_5_4 13CO_6_5
13CO_7_6 13CO_8_7 13CO_9_8 C180_1_0 C180_2_1 C180_3_2 C180_4_3 C180_5_4 C180_6_5 C180_7_6
C180_8_7 C180_9_8 CS_1_0 CS_2_1 CS_3_2 CS_4_3 CS_5_4 CS_6_5 CS_7_6 CS_8_7 CS_9_8 CS_10_9
CS_11_10 CS_12_11 CS_13_12 CS_14_13 CS_15_14 CS_16_15 CS_17_16 CS_18_17 CS_19_18 CS_12_19
SiO_1_0 SiO_2_1 SiO_3_2 SiO_4_3 SiO_5_4 SiO_6_5 SiO_7_6 SiO_8_7 SiO_9_8 SiO_10_9 SiO_11_10
SiO_12_11 SiO_13_12 SiO_14_13 SiO_15_14 SiO_16_15 SiO_17_16 SiO_18_17 SiO_19_18 SiO_20_19
SiO_21_20 SiO_22_21 SiO_23_22'

```

3.3.1.11 Shifting the Phase Center

The plot's phase center can be shifted in the **Plots > Trans** tab. Enter the X and Y shifts in units of arcseconds in the blank boxes under **Phase center shift**.

3.3.1.12 Plot Ranges

The X and Y ranges of the plot can be set in the **Plots > Axes** tab. By default, the circle next to **Automatic** will be checked, and the ranges will be auto-scaled. To define the range, click on the circle below **Automatic** and enter a minimum and maximum value in the blank boxes (as for the X Axis in Figure 3.2. Note that if identical values are placed in the blank boxes ($xmin=xmax$ and/or $ymin=ymax$), then the values will be ignored and a best guess will be made to auto-range that axis.

3.3.1.13 Plot Labels

The plot and axes labels which are displayed in the plot window are set in the **Plots > Canvas** tab. To change the plot title, under **Canvas Title**, click on the circle next to the blank box and enter the desired text. To change the X- and Y-axis labels, similarly click on the circles next to the blank boxes under **Show X Axis** and **Show Y Axis** and type the desired text in the blank box. To display these new labels, simply click the **Plot** button.

The user can determine the locations of axis labels in the **Plots > Axes** tab. The X-axis label switches from the bottom to the top of the plot depending on what is selected for **Attach to:**. Similarly for the Y-Axis, the user can choose to attach axis labels and tick marks to the **Top** or **Bottom** (note that the axis labels have been attached to the **Bottom** and **Right** in Figure 3.2.

Finally, axis labels can be removed all together by unchecking the boxes next to **Show X Axis** and **Show Y Axis** on the **Plots > Canvas** tab.

3.3.1.14 Grid Lines

A grid of lines can be superimposed on the plot using **Grid Lines** in the **Plots > Canvas** tab. “Major” grid lines are drawn at the locations of major tick marks, while “minor” grid lines are drawn at minor tick marks.

Grid line colors, thicknesses, and styles are selected independently for the “major” and “minor” grid lines. Desired line thickness should be typed into the blank boxes just to the right of the **Major** and **Minor** labels. Colors are set by clicking on the ... buttons. The blank boxes to the left of the ... buttons will then contain the hex codes for the selected colors (e.g., “808080”). Line styles can also be selected from the drop-down menus to the right of ... buttons.

3.3.1.15 Legend

A plot symbol legend can be added to the plot by clicking on the checkbox next to **Legend** in the **Plots > Canvas** tab. However, given the current functionalities of **plotms**, a symbol legend is of very limited use. This option is useful when overplotting data.

3.3.1.16 The Options Tab

A few miscellaneous options are available in the **Options** tab, the last tab at the top of the **plotms** window. The **Tool Button Style** drop-down menu determines if icons and/or text represent the buttons in the toolbar near the bottom of the **plotms** window.

The **Log Events** drop down menu determines how verbose **plotms** is in documenting its actions on the command line.

There is a checkbox that determines the persistence of regions and annotations on new plots, labelled **When changing plot axes, clear any existing regions and annotations**.

A useful option is the **fixed size for cached image** checkbox. It determines how large the dots in the panel are with respect to the screen resolution. The values influence how the data is redrawn on the panel. When the **Screen resolution** is selected, the **plotms** window can be resized without redrawing on the canvas – a considerable speedup for large data sets. The penalty is that the dots of the data points are the size of a pixel on the screen, which may be very small for high resolution monitors.

Finally, the **File chooser history limit** determines the number of remembered directories in the file loading pop-up of the **Browse** selection of the **Data** tab.

3.3.1.17 Iteration

In many cases, it is desirable to iterate through the data that were selected in the *Data* tab. A typical example is to display a single baseline in a time vs. amplitude plot and then proceed to the next baselines step by step. This can be done via the *Page* tab on the left hand side of **plotms**.

A drop-down menu allows you to select the parameter to be iterated on, such as baseline or spw (press *plot* after changing your selection). The plot titles in the main panel in `plotms` show which data slice is currently displayed. To proceed to the next plot use the *green buttons* below the main panel. The different button symbols let you to proceed panel by panel or to jump to the first or last panel directly. The number of plots per page can be selected under *Options- \dot{z} Grid*, the last of the top row of tabs (corresponding to the *gridrows* and *gridcols* in the command line interface).

There are two scaling options for the axes: *Global* and *Self*. *Global* will use a common axis range based on data loaded with the selection criteria specified in the *Data* tab. *Self* readjusts the axes scaling to the data for each individual panel of the iteration.

Below, one can invoke multiple panels per display by selecting the number of rows and columns to be displayed on the canvas.

3.3.1.18 Overplotting

Different values of the same dataset can be shown at the same time. E.g. to add a second y-axis, press the “Add Y Axis Data” button under the “Axes” tab. Then select the parameters for the newly created axis by selecting from the now available “Y Axis Data” drop-down menu. If the two y-axes have the same units, they can be displayed both on the same axis. If they are different, e.g. Amplitude and Elevation (both versus time; see Fig. 3.4), one axis should be attached to the left and the other to the right hand side of the plot. Using more than a single y-axis data is also reflected in the “Display” tab where a drop-down menu appears in order to select multiple y-axis options.

In the `plotms` input interface, you can overplot by invoking `plotms` more than once with *clearplot=F*. Each run of `plotms` corresponds to a plot to go on top of previous ones.

3.3.1.19 Plotting Multiple Data Sets

`plotms` can also plot more than a single dataset in separate panels. To do so, press “Add Plot” next to the “Plot” button. This will bring up a new data window where the plot parameters are defined. Right-click options are used to “Minimize”, “Maximize”, or “Close” these panels which helps to keep a better overview on the individual datasets. If *Options- \dot{z} Grid* is selected to have more than a single panel, the different datasets will be shown side by side.

When `plotms` is run from the command line, the location of the plots can be defined as follows. *gridcols* and *gridrows* define the number of plots on the screen. To define the location where a subplot is to appear on this grid, use *colindex* and *rowindex*. If one uses a *plotindex*, this will be used as a label to address the plot. Each call of `plotms` with the same *plotindex* will overplot on the subplot where *plotindex* was defined the first time. Here is an example on multiple `plotms` calls:

```
#Plot in the second column, first row of a 2x2 grid and define this plot as plotindex=0
plotms(vis='vis1.ms', gridrows=2, gridcols=2, colindex=1, rowindex=0)
```

```
#Overplot in the same panel using a different axis and symbol for the second plot.
```

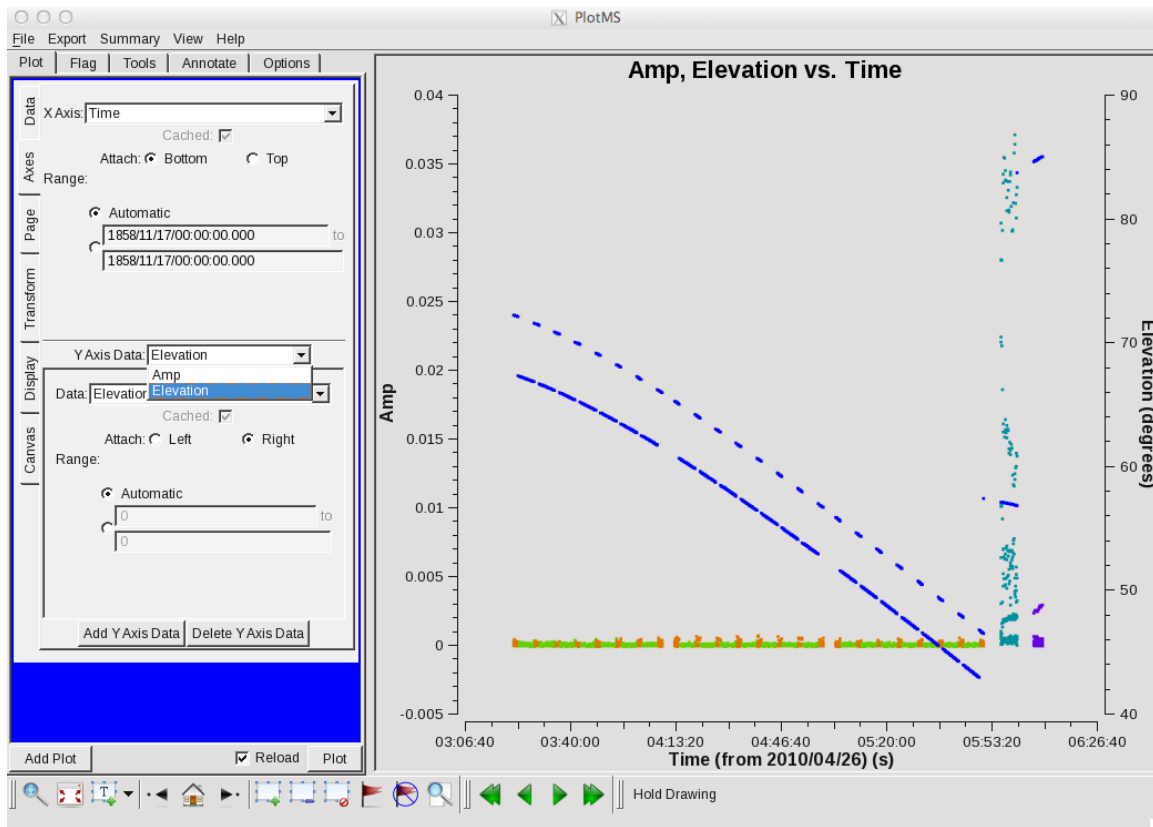


Figure 3.4: Overplotting in plotms: Two different y-axes have been chosen for this plot, amplitude and elevation.

```

plotms(vis='vis2.ms', clearplots=False, plotindex=1, rowindex=0,
colindex=1, gridrows=2, gridcols=2, yaxislocation='right', symbolshape='circle')

#Define a second plot and give it a label plotindex=2, in the lower right corner of the grid.
plotms(vis='vis1.ms', clearplots=False, plotindex=2, rowindex=1,colindex=1, gridrows=2, gridcols=2)

#Move the plot with the overplot one panel to the left. This requires clearing
#the plots and rerunning the script specifications with the new plot locations.
plotms(vis='vis1.ms', gridrows=2, gridcols=2, colindex=0, rowindex=0, symbolshape='diamond')
plotms(vis='vis2.ms', clearplots=False, plotindex=1, rowindex=0, colindex=0,
      gridrows=2, gridcols=2,yaxislocation='right',
      symbolshape='circle')
plotms(vis=self.ms, clearplots=False, plotindex=2, rowindex=1,colindex=1,gridrows=2, gridcols=2)

```

3.3.1.20 Saving your plot

You can save a copy of a plot to file in the **Plots > Export** tab. Click the **Browse** button for a GUI-based selection of the directory and file name to which the plot will be saved. The file format can also be determined in this GUI by the suffix given to the filename: *.png* (PNG), *.jpg* (JPG), *.ps* (PS), *.pdf* (PDF), and *.txt* (TEXT). Alternatively, the file format can be selected from the **Format** drop-down menu located just below the **Browse** button. In this case, `plotms` will add a suffix to the file name depending on the format chosen.

ALERT: The plot files produced by the PS and PDF options can be large and time-consuming to export. The JPG is the smallest.

The exported plot resolution can be manipulated using the **High Resolution**, **DPI**, and **Size** options.

Click on **Export** to create the file, you may select to either plot only the current page or all pages (filenames will be automatically incremented).

the *TEXT* format will not save an image but the data points themselves. This allows one to dump the current plot into a file that is used in other programs for further processing. The reported data is the same as when using the *locate* button in `plotms` and the format looks like:

```
# x y chan scan field ant1 ant2 ant1name ant2name time freq spw corr offset currchunk irel
# Real Imag None None None None None None None MJD(seconds) GHz None None None None None
0.282938 0.0387583 31 5 2 1 12 ea02@E02 ea21@E01 4778968956.000 36.308479452 1 RR 26 0 26
0.263241 -0.00806698 31 7 2 1 12 ea02@E02 ea21@E01 4778969356.000 36.308479452 1 RR 29 1 28
0.258207 0.0301206 31 9 2 1 12 ea02@E02 ea21@E01 4778969745.000 36.308479452 1 RR 30 2 28
0.311155 -0.0180511 31 11 2 1 12 ea02@E02 ea21@E01 4778970133.250 36.308479452 1 RR 31 3 28
0.284589 -0.0628808 31 13 2 1 12 ea02@E02 ea21@E01 4778970522.250 36.308479452 1 RR 32 4 28
```

where *x* and *y* are the two plotted axes and the other columns contain additional information such as the baselines or frequencies. The three last columns *offset*, *corrchunk*, and *irel* are internal data management items for `plotms` and you most likely will never use them.

3.3.1.21 Exiting plotms

To exit the `plotms` GUI, select **Quit** from the **File** menu at the top of the `plotms` window. You can also dismiss the window by killing it with the “X” on the frame.

Alternatively, you can just leave it alone, and `plotms` will keep running in the background. If the data file changes in the background, you can force reloading the data via the ‘force reload’ checkbox next to the ‘Plot’ button. Alternatively, press SHIFT while clicking on ‘Plot’ for the same purpose.

3.3.2 Plotting and Editing using plotxy

Inside the Toolkit: Access to *matplotlib* is also provided through the `pl` tool. See below for a description of the `pl` tool functions.

ALERT: The `plotxy` code is fragile and slow, and is being replaced by the `plotms` (§ 3.3.1). We retain `plotxy` in this release as not all functionality is available yet in `plotms`.

`Plotxy` is a tool for visualizing and editing visibility data. Unlike `plotms`, it is useful in scripting, as it can non-interactively produce a hardcopy plot (see § 3.3.2.13). It also has multi-plot (§ 3.3.2.8), iteration (§ 3.3.2.3), and overplotting (§ 3.3.2.4) functionality—unlike `plotms` in the current release. `Plotxy` uses the `matplotlib` plotting library to display its plots. You can find information on `matplotlib` at <http://matplotlib.sourceforge.net/>.

To bring up this plotter use the `plotxy` task. The inputs are:

```
# plotxy :: X-Y plotter/interactive flagger for visibility data

vis           =      ''      # Name of input visibility
xaxis         =      'time'   # X-axis: def = 'time': see help for options
yaxis         =      'amp'    # Y-axis: def = 'amp': see help for options
  datacolumn  =      'data'   # data (raw), corrected, model, residual (corrected - model)

selectdata    =      False   # Other data selection parameters
spw           =      ''      # spectral window:channels: ''==>all, spw='1:5~57'
field         =      ''      # field names or index of calibrators: ''==>all
averagemode   =      ''      # Select averaging type: 'vector', 'scalar'
restfreq      =      ''      # a frequency quanta or transition name. see help for options
extendflag    =      False   # Have flagging extend to other data points?
subplot       =      111     # Panel number on display screen (yxn)
plotsymbol    =      '.'     # Options include . : , o ^ v > < s + x D d 2 3 4 h H | _
plotcolor     =      'darkcyn' # Plot color
plotrange     =      [-1, -1, -1, -1] # The range of data to be plotted (see help)
multicolor    =      'corr'   # Plot in different colors: Options: none, both, chan, corr
selectplot    =      False   # Select additional plotting options (e.g, fontsize, title,etc)
overplot      =      False   # Overplot on current plot (if possible)
showflags     =      False   # Show flagged data?
interactive    =      True    # Show plot on gui?
figfile       =      ''      # ''= no plot hardcopy, otherwise supply name
```

ALERT: The `plotxy` task expects all of the scratch columns to be present in the MS, even if it is not asked to plot the contents. If you get an error to the effect *"Invalid Table operation: Table: cannot add a column"* then use `clearcal()` to force these columns to be made in the MS. Note that this will clear anything in all scratch columns (in case some were actually there and being used).

Setting `selectdata=True` opens up the selection sub-parameters:

```
selectdata    =      True    # Other data selection parameters
  antenna     =      ''      # antenna/baselines: ''==>all, antenna = '3,VA04'
```

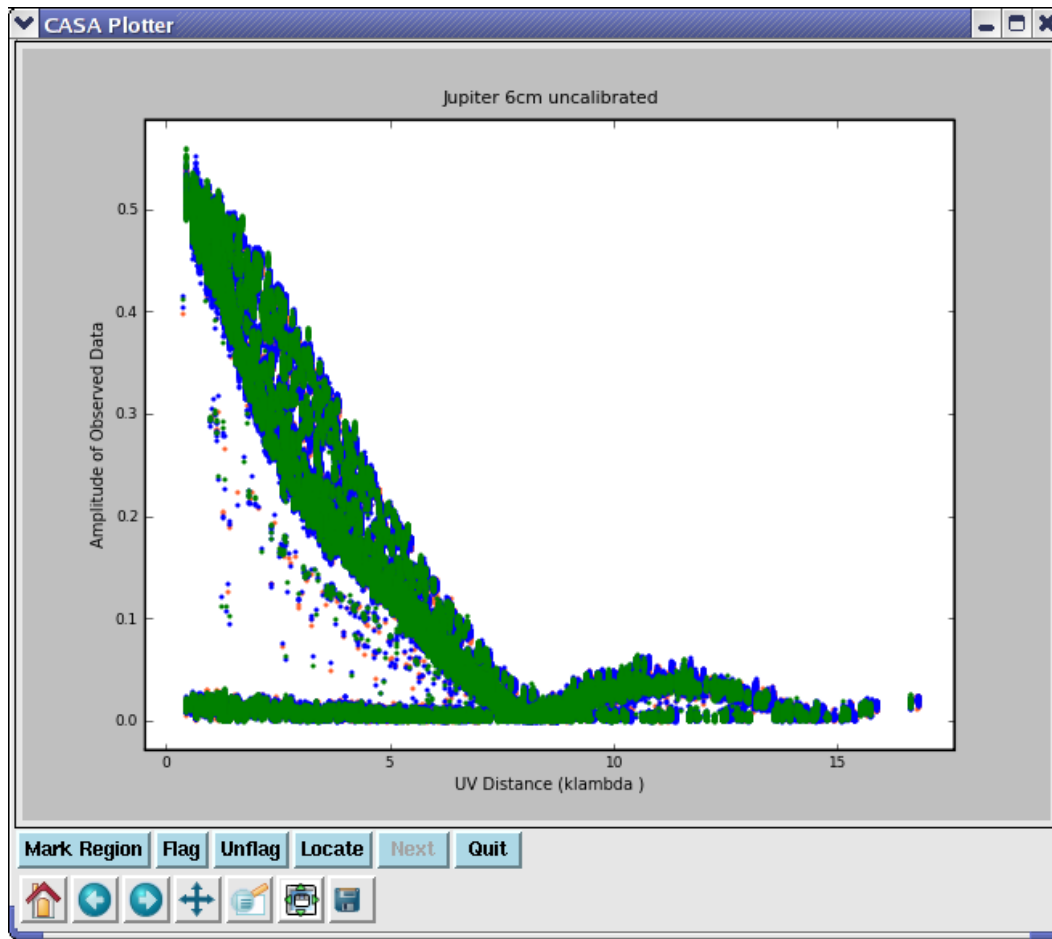


Figure 3.5: The plotxy plotter, showing the Jupiter data versus uv-distance. You can see bad data in this plot. The **bottom set of buttons** on the lower left are: 1,2,3) **Home, Back, and Forward**. Click to navigate between previously defined views (akin to web navigation). 4) **Pan**. Click and drag to pan to a new position. 5) **Zoom**. Click to define a rectangular region for zooming. 6) **Subplot Configuration**. Click to configure the parameters of the subplot and spaces for the figures. 7) **Save**. Click to launch a file save dialog box. The **upper set of buttons in the lower left** are: 1) **Mark Region**. Press this to begin marking regions (rather than zooming or panning). 2,3,4) **Flag, Unflag, Locate**. Click on these to flag, unflag, or list the data within the marked regions. 5) **Next**. Click to move to the next in a series of iterated plots. Finally, the **cursor readout** is on the bottom right.

```

timerange = '' # time range: ''=>all
correlation = '' # correlations: default = ''
scan = '' # scan numbers: Not yet implemented
feed = '' # multi-feed numbers: Not yet implemented
array = '' # array numbers: Not yet implemented
uvrange = '' # uv range''=>all; uvrange = '0~100kl' (default unit=meters)

```


These are described in § 2.3.

Averaging is controlled with the set of parameters

```

averagemode      = 'vector' # Select averaging type: vector, scalar
timebin          = '0'     # Length of time-interval in seconds to average
crossscans       = False   # Have time averaging cross scan boundaries?
crossbls         = False   # have averaging cross over baselines?
crossarrays      = False   # have averaging cross over arrays?
stackspw         = False   # stack multiple spw on top of each other?
width            = '1'     # Number of channels to average

```

See § 3.3.2.9 below for more on averaging.

You can extend the flagging beyond the data cell plotted:

```

extendflag       = True    # Have flagging extend to other data points?
extendcorr       = ''      # flagging correlation extension type
extendchan       = ''      # flagging channel extension type
extendspw        = ''      # flagging spectral window extension type
extendant        = ''      # flagging antenna extension type
extendtime       = ''      # flagging time extension type

```

See § 3.3.2.11 below for more on flag extension.

The *restfreq* parameter can be set to a transition or frequency:

```

restfreq         = 'HI'    # a frequency quanta or transition name. see help for options
frame            = 'LSRK'  # frequency frame for spectral axis. see help for options
doppler          = 'RADIO' # doppler mode. see help for options

```

See § 3.3.2.12 below for more on setting rest frequencies and frames.

Setting *selectplot=True* will open up a set of plotting control sub-parameters. These are described in § 3.3.2.2 below.

The *interactive* and *figfile* parameters allow non-interactive production of hardcopy plots. See § 3.3.2.13 for more details on saving plots to disk.

The *iteration*, *overplot*, *plotrange*, *plotsymbol*, *showflags* and *subplot* parameters deserve extra explanation, and are described below.

For example:

```

plotxy(vis='jupiter6cm.ms', # jupiter 6cm dataset
        xaxis='uvdist',     # plot uv-distance on x-axis
        yaxis='amp',        # plot amplitude on y-axis
        field='JUPITER',    # plot only JUPITER
        selectdata=True,    # open data selection
        correlation='RR,LL', # plot RR and LL correlations
        selectplot=True,    # open plot controls
        title = 'Jupiter 6cm uncalibrated') # give it a title

```

The plotter resulting from these settings is shown in figure 3.5.

ALERT: The `plotxy` task still has a number of issues. The averaging has been greatly speeded up in this release, but there are cases where the plots will be made incorrectly. In particular, there are problems plotting multiple `spw` at the same time. There are sometimes also cases where data that you have flagged in `plotxy` from averaged data is done so incorrectly. This task is under active development for the next cycle to fix these remaining problems, so users should be aware of this.

ALERT: Another known problem with (`plotxy`) is that it fails if the path to your working directory contains spaces in its name, e.g. `/users/smyers/MyTest/` is fine, but `/users/smyers/My Test/` is not!

3.3.2.1 GUI Plot Control

You can use the various buttons on the `plotxy` GUI to control its operation – in particular, to determine flagging and unflagging behaviors.

There is a standard row of buttons at the bottom. These include (left to right):

- **Home** — The “house” button (1st on left) returns to the original zoom level.
- **Step** — The left and right arrow buttons (2nd and 3rd from left) step through the zoom settings you’ve visited.
- **Pan** — The “four-arrow button” (4th from left) lets you pan in zoomed plot.
- **Zoom** — The most useful is the “magnifying glass” (5th from the left) which lets you draw a box and zoom in on the plot.
- **Panels** — The “window-thingy” button (second from right) brings up a menu to adjust the panel placement in the plot.
- **Save** — The “disk” button (last on right) saves a `.png` copy of the plot to a generically named file on disk.

In a row above these, there are a set of other buttons (left to right):

- **Mark Region** — If depressed lets you draw rectangles to mark points in the panels. This is done by left-clicking and dragging the mouse. You can Mark multiple boxes before doing something. Clicking the button again will un-depress it and forget the regions. ESC will remove the last region marked.
- **Flag** — Click this to Flag the points in a marked region.
- **Unflag** — Click this to Unflag any flagged point that would be in that region (even if invisible).
- **Locate** — Print out some information to the logger on points in the marked regions.

- **Next** — Step to the next plot in an iteration.
- **Quit** — Exit `plotcal`, clear the window and detach from the MS.

These buttons are shared with the `plotcal` tool.

3.3.2.2 The `selectplot` Parameters

These parameters work in concert with the native `matplotlib` functionality to enable flexible representations of data displays.

Setting `selectplot=True` will open up a set of plotting control sub-parameters:

```
selectplot      =      True   # Select additional plotting options (e.g, fontsize, title,etc)
  markersize    =      5.0   # Size of plotted marks
  linewidth     =      1.0   # Width of plotted lines
  skipnrows     =        1   # Plot every nth point
  newplot       =      False  # Replace the last plot or not when overplotting
  clearpanel    =      'Auto' # Specify if old plots are cleared or not
  title         =          '' # Plot title (above plot)
  xlabel        =          '' # Label for x-axis
  ylabel        =          '' # Label for y-axis
  fontsize      =     10.0   # Font size for labels
  windowsize    =      5.0   # Window size: not yet implemented
```

The `markersize` parameter will change the size of the plot symbols. Increasing it will help legibility when doing screen shots. Decreasing it can help in congested plots. The `linewidth` parameter will do similar things to the lines.

The `skipnrows` parameter, if set to an integer n greater than 1, will allow only every n th point to be plotted. It does this, as the name suggests, by skipping over whole rows of the MS, so beware (channels are all within the same row for a given `spw`). Be careful flagging on data where you have skipped points! Note that you can also reduce the number of points plotted via averaging (§ 3.3.2.9) or channel striding in the `spw` specification (§ 2.3.3).

The `newplot` toggle lets you choose whether or not the last layer plotted is replaced when `overplot=True`, or whether a new layer is added.

The `clearpanel` parameter turns on/off the clearing of plot panels that lie under the current panel layer being plotted. The options are: `'none'` (clear nothing), `'auto'` (automatically clear the plotting area), `'current'` (clear the current plot area only), and `'all'` (clear the whole plot panel).

The `title`, `xlabels`, and `ylabels` parameters can be used to change the plot title and axes labels.

The `fontsize` parameter is useful in order to enlarge the label fonts so as to be visible when making plots for screen capture, or just to improve legibility. Shrinking can help if you have lots of panels on the plot also.

Inside the Toolkit:

For even more functionality, you can access the `pl` tool directly using `Py-lab` functions that allow one to annotate, alter, or add to any plot displayed in the `matplotlib` plotter (e.g. `plotxy`).

The *window* parameter is supposed to allow adjustments on the window size. **ALERT:** This currently does nothing, unless you set it below 1.0, in which case it will produce an error.

3.3.2.3 The *iteration* parameter

Under *Page* one can select plot iterations parameters. There are currently four iteration options available: *'field'*, *'antenna'*, and *'baseline'*. If one of these options is chosen, the data will be split into separate plot displays for each value of the iteration axis (e.g., for the VLA, the *'antenna'* option will get you 27 displays, one for each antenna). An example use of iteration:

```
# choose channel averaging, every 5 channels
plotxy('n5921.ms', 'channel', subplot=221, iteration='antenna', width='5')
```

The results of this are shown in Figure 3.6. Note that this example combines the use of *width*, *iteration* and *subplot*.

NOTE: If you use *iteration='antenna'* or *'baseline'*, be aware if you have set *antenna* selection. You can also control whether you see auto-correlations or not using the appropriate syntax, e.g. *antenna='*ℓℓ*'* or *antenna='*ℓℓℓ'* (§ 2.3.4.1).

3.3.2.4 The *overplot* parameter

The *overplot* parameter toggles whether the current plot will be overlaid on the previous plot or subpanel (via the *subplot* setting, § section:edit.plot.plotxy.subplot) or will overwrite it. The default is *False* and the new plot will replace the old.

The *overplot* parameter interacts with the *newplot* sub-parameter (see § 3.3.2.2).

See § 3.3.2.7 for an example using *overplot*.

3.3.2.5 The *plorange* parameter

The *plorange* parameter can be used to specify the size of the plot. The format is [*xmin*, *xmax*, *ymin*, *ymax*]. The units are those on the plot. For example,

```
plorange = [-20,100,15,30]
```

Note that if *xmin=xmax* and/or *ymin=ymax*, then the values will be ignored and a best guess will be made to auto-range that axis.

Unfortunately, the units for the time axis must be in Julian seconds. This is somewhat inconvenient as the usual time parameter is given in Julian days. To calculate the Julian seconds the *me.epoch* tool can be used. An example: For 02:00 UT on 2012/05/22, the value of MJD seconds can be calculated via

```
86400*(me.epoch('utc', '2012/05/22')['m0']['value']+2/24.)
```

which results in *4844368800.0*.

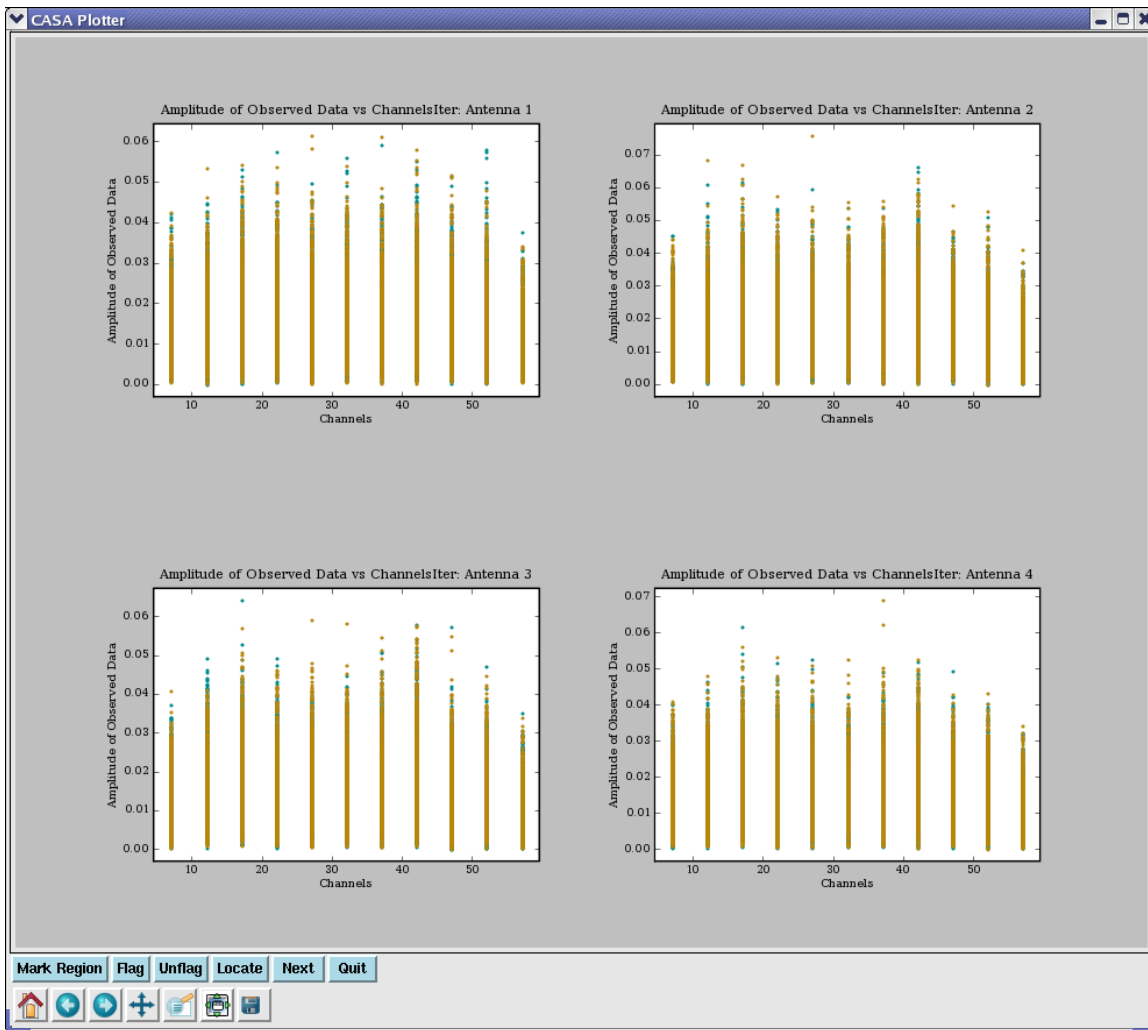


Figure 3.6: The plotxy iteration plot. The first set of plots from the example in § 3.3.2.3 with `iteration='antenna'`. Each time you press the **Next** button, you get the next series of plots.

3.3.2.6 The `plotsymbol` parameter

The `plotsymbol` parameter defines both the line or symbol for the data being drawn as well as the color; from the matplotlib online documentation (e.g., type `pl.plot?` for help):

The following line styles are supported:

- : solid line
- : dashed line
- .: dash-dot line
- : : dotted line
- . : points

```

,      : pixels
o      : circle symbols
^      : triangle up symbols
v      : triangle down symbols
<      : triangle left symbols
>      : triangle right symbols
s      : square symbols
+      : plus symbols
x      : cross symbols
D      : diamond symbols
d      : thin diamond symbols
1      : tripod down symbols
2      : tripod up symbols
3      : tripod left symbols
4      : tripod right symbols
h      : hexagon symbols
H      : rotated hexagon symbols
p      : pentagon symbols
|      : vertical line symbols
_      : horizontal line symbols
steps : use gnuplot style 'steps' # kwarg only

```

The following color abbreviations are supported

```

b : blue
g : green
r : red
c : cyan
m : magenta
y : yellow
k : black
w : white

```

In addition, you can specify colors in many weird and wonderful ways, including full names 'green', hex strings '#008000', RGB or RGBA tuples (0,1,0,1) or grayscale intensities as a string '0.8'.

Line styles and colors are combined in a single format string, as in 'bo' for blue circles.

3.3.2.7 The *showflags* parameter

The *showflags* parameter determines whether *only* unflagged data (*showflags=False*) or flagged (*showflags=True*) data is plotted by this execution. The default is *False* and will show only unflagged "good" data.

Note that if you want to plot both unflagged and flagged data, in different colors, then you need to run `plotxy` twice using *overplot* (see § 3.3.2.4) the second time, e.g.

```

> plotxy(vis="myfile", xaxis='uvdist', yaxis='amp' )
> plotxy(vis="myfile", xaxis='uvdist', yaxis='amp', overplot=True, showflags=True )

```

3.3.2.8 The *subplot* parameter

The *subplot* parameter takes three numbers. The first is the number of y panels (stacking vertically), the second is the number of xpanels (stacking horizontally) and the third is the number of the panel you want to draw into. For example, *subplot=212* would draw into the lower of two panels stacked vertically in the figure.

An example use of subplot capability is shown in Fig 3.7. These were drawn with the commands (for the top, bottom left, and bottom right panels respectively):

```
plotxy('n5921.ms','channel',      # plot channels for the n5921.ms data set
       field='0',                 # plot only first field
       datacolumn='corrected',    # plot corrected data
       plotcolor='',             # over-ride default plot color
       plotsymbol='go',          # use green circles
       subplot=211)              # plot to the top of two panels

plotxy('n5921.ms','x',           # plot antennas for n5921.ms data set
       field='0',                 # plot only first field
       datacolumn='corrected',    # plot corrected data
       subplot=223,              # plot to 3rd panel (lower left) in 2x2 grid
       plotcolor='',             # over-ride default plot color
       plotsymbol='r.')          # red dots

plotxy('n5921.ms','u','v',      # plot uv-coverage for n5921.ms data set
       field='0',                 # plot only first field
       datacolumn='corrected',    # plot corrected data
       subplot=224,              # plot to the lower right in a 2x2 grid
       plotcolor='',             # over-ride default plot color
       plotsymbol='b,')          # blue, somewhat larger dots
                                  # NOTE: You can change the gridding
                                  # and panel size by manipulating
                                  # the ny x nx grid.
```

See also § 3.3.2.3 above, and Figure 3.6 for an example of channel averaging using *iteration* and *subplot*.

3.3.2.9 Averaging in *plotxy*

The averaging parameters and sub-parameters are:

```
averagemode    = 'vector' # Select averaging type: vector, scalar
timebin        = '0' # length of time in seconds to average, default='0', or: 'all'
crossscans     = False # have time averaging cross over scans?
crossbls       = False # have averaging cross over baselines?
crossarrays    = False # have averaging cross over arrays?
stackspw       = False # stack multiple spw on top of each other?
width          = '1' # number of channels to average, default: '1', or: 'all', 'allspw'
```

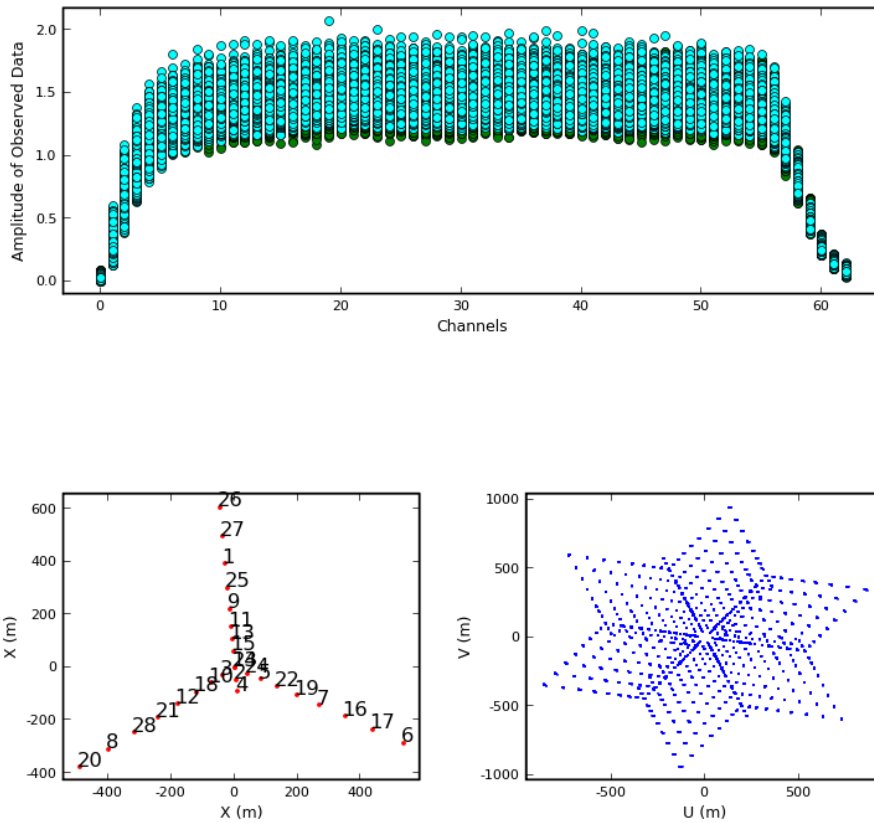


Figure 3.7: Multi-panel display of visibility versus channel (**top**), antenna array configuration (**bottom left**) and the resulting uv coverage (**bottom right**). The commands to make these three panels respectively are: 1) `plotxy('ngc5921.ms', xaxis='channel', datacolumn='data', field='0', subplot=211, plotcolor='', plotsymbol='go')` 2) `plotxy('ngc5921.ms', xaxis='x', field='0', subplot=223, plotsymbol='r.')` 3) `plotxy('ngc5921.ms', xaxis='u', yaxis='v', field='0', subplot=224, plotsymbol='b', figfile='ngc5921_multiplot.png')`.

The choice of *averagemode* controls how the amplitudes are calculated in the average. The default mode is *vector*, where the complex average is formed by averaging the real and imaginary parts of the relevant visibilities. If *scalar* is chosen, then the amplitude of the average is formed by a scalar average of the individual visibility amplitudes.

Time averaging is effected by setting the *timebin* parameter to a value larger than the integration time. Currently, *timebin* takes a string containing the averaging time in seconds, e.g.

```
timebin = '60.0'
```


to plot one-minute averages.

Channel averaging is invoked by setting *width* to a value greater than 1. Currently, the averaging *width* is given as a number of channels.

By default, the averaging will not cross *scan* boundaries (as set in the import process). However, if *crossscans=True*, then averaging will cross scans.

Note that data taken in different sub-arrays are never averaged together. Likewise, there is no way to plot data averaged over *field*.

3.3.2.10 Interactive Flagging in plotxy

Interactive flagging, on the principle of “see it — flag it”, is possible on the X-Y display of the data plotted by `plotxy`. The user can use the cursor to mark one or more regions, and then flag, unflag, or list the data that falls in these zones of the display.

Hint!

In the plotting environments such as `plotxy`, the *ESC* key can be used to remove the last region box drawn.

There is a row of buttons below the plot in the window. You can punch the **Mark Region** button (which will appear to depress), then mark a region by left-clicking and dragging the mouse (each click and drag will mark an additional region). You can get rid of all your regions by clicking again on the **Mark Region** button (which will appear to un-depress), or you can use the *ESC* key to remove the last box you drew. Once regions are marked, you can then click on one of the other buttons to take action:

1. **Flag** — flag the points in the region(s),
2. **Unflag** — unflag flagged points in the region(s),
3. **Locate** — spew out a list of the points in the region(s) to the logger (Warning: this could be a long list!).

Whenever you click on a button, that action occurs without forcing a disk-write (unlike previous versions). If you quit `plotxy` and re-enter, you will see your previous edits.

A table with the name *jmsname.j.flagversions* (where *vis=jmsname.j*) will be created in the same directory if it does not exist already.

It is recommended that you save important flagging stages using the *flagmanager* task (§ 3.2).

3.3.2.11 Flag extension in plotxy

Flag extension is controlled using *extendflag=T* and its sub-parameters:

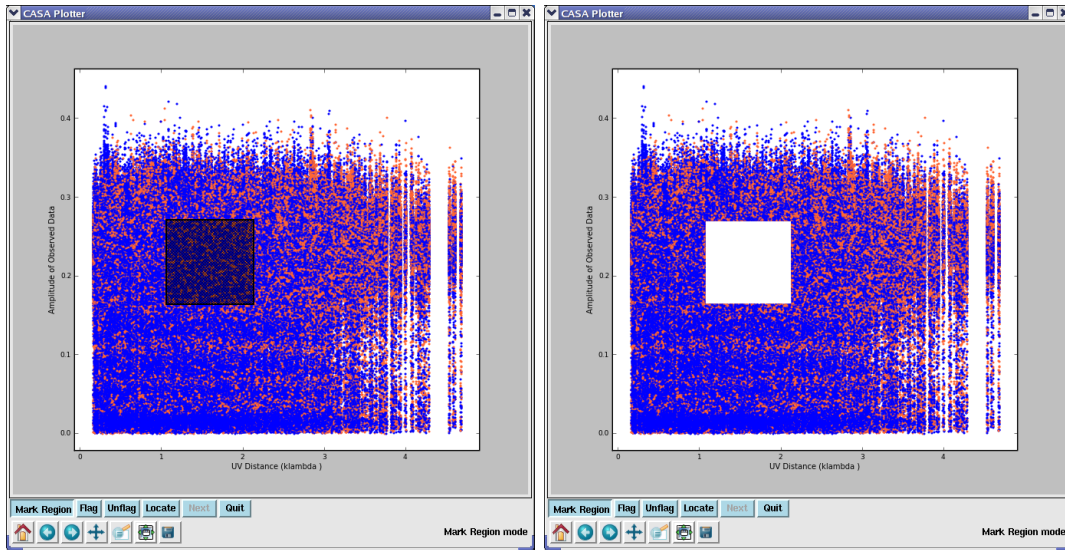


Figure 3.8: Plot of amplitude versus uv distance, before (left) and after (right) flagging two marked regions. The call was: `plotxy(vis='ngc5921.ms', xaxis='uvdist', field='1445*')`.

```

extendflag      = True      # Have flagging extend to other data points?
  extendcorr    = ''        # flagging correlation extension type
  extendchan    = ''        # flagging channel extension type
  extendspw     = ''        # flagging spectral window extension type
  extendant     = ''        # flagging antenna extension type
  extendtime    = ''        # flagging time extension type

```

The use of `extendflag` enables the user to plot a subset of the data and extend the flagging to a wider set.

ALERT: Using the `extendflag` options will greatly slow down the flagging in `plotxy`. You will see a long delay after hitting the **Flag** button, with lots of logger messages as it goes through each flag. Fixing this requires a refactoring of `plotxy` which is underway starting in Patch 4 development.

Setting `extendchan='all'` will extend the flagging to other channels in the same `spw` as the displayed point. For example, if `spw='0:0'` and channel 0 is displayed, then flagging will extend to all channels in `spw 0`.

The `extendcorr` sub-parameter will extend the flagging beyond the correlations displayed. If `extendcorr='all'`, then all correlations will be flagged, e.g. with RR displayed RR,RL,LR,LL will be flagged. If `extendcorr='half'`, then the extension will be to those correlations in common with that show, e.g. with RR displayed then RR,RL,LR will be flagged.

Setting `extendspw='all'` will extend the flagging to all other `spw` for the selection. Using the same example as above, with `spw='0:0'` displayed, then channel 0 in ALL `spw` will be flagged. Note that use of `extendspw` could result in unintended behavior if the `spw` have different numbers of channels, or if it is used in conjunction with `extendchan`.

WARNING: use of the following options, particularly in conjunction with other flag extensions, may lead to deletion of much more data than desired. Be careful!

Setting `extendant='all'` will extend the flagging to all baselines that have antennas in common with those displayed and marked. For example, if `antenna='1&2'` is shown, then ALL baselines to BOTH antennas 1 and 2 will be flagged. Currently, there is no option to extend the flag to ONLY baselines to the first (or second) antenna in a displayed pair.

Setting `extendtime='all'` will extend the flagging to all times matching the other selection or extension for the data in the marked region.

3.3.2.12 Setting rest frequencies in plotxy

The `restfreq` parameter can be set to a transition or frequency and expands to allow setting of frame information. For example,

```
restfreq      = 'HI'      # a frequency quanta or transition name. see help for options
frame         = 'LSRK'    # frequency frame for spectral axis. see help for options
doppler       = 'RADIO'   # doppler mode. see help for options
```

Examples of transitions include:

```
restfreq='1420405751.786Hz' # 21cm HI frequency
restfreq='HI'                # 21cm HI transition name
restfreq='115.2712GHz'       # CO 1-0 line frequency
```

For a list of known lines in the CASA *measures* system, use the toolkit command `me.linelist()`. For example:

```
CASA <14>: me.linelist()
Out[14]: 'C109A CI CII166A DI H107A H110A H138B H166A H240A H272A
H2CO HE110A HE138B HI OH1612 OH1665 OH1667 OH1720'
```

ALERT: The list of known lines in CASA is currently very restricted, and will be increased in upcoming releases (to include lines in ALMA bands for example).

You can use the `me.spectralline` tool method to turn transition names into frequencies

```
CASA <16>: me.spectralline('HI')
Out[17]:
{'m0': {'unit': 'Hz', 'value': 1420405751.786},
 'refer': 'REST',
 'type': 'frequency'}
```

(not necessary for this task, but possibly useful).

The `frame` sub-parameter sets the frequency frame. The allowed options can be listed using the `me.listcodes` method on the `me.frequency()` method, e.g.

```
CASA <17>: me.listcodes(me.frequency())
Out [17]:
{'extra': array([],
  dtype='|S1'),
 'normal': array(['REST', 'LSRK', 'LSRD', 'BARY', 'GEO', 'TOPO', 'GALACTO', 'LGROUP',
  'CMB'],
  dtype='|S8')}
```

The *doppler* sub-parameter likewise sets the Doppler system. The allowed codes can be listed using the *me.listcodes* method on the *me.doppler()* method,

```
CASA <18>: me.listcodes(me.doppler())
Out [18]:
{'extra': array([],
  dtype='|S1'),
 'normal': array(['RADIO', 'Z', 'RATIO', 'BETA', 'GAMMA', 'OPTICAL', 'TRUE',
  'RELATIVISTIC'],
  dtype='|S13')}
```

For most cases the *'RADIO'* Doppler system is appropriate, but be aware of differences.

For more information on frequency frames and spectral coordinate systems, see the paper by Greisen et al. (A&A, 446, 747, 2006) ².

3.3.2.13 Printing from plotxy

There are two ways to get hardcopy plots in *plotxy*.

The first is to use the “disk save” icon from the interactive plot GUI to print the current plot. This will bring up a sub-menu GUI that will allow you to choose the filename and format. The allowed formats are *.png* (PNG), *.eps* (EPS), and *svg* (SVG). If you give the filename with a suffix (*.png*, *.eps*, or *svg*) it will make a plot of that type. Otherwise it will put a suffix on depending on the format chosen from the menu.

ALERT: The plot files produced by the EPS option can be large, and the SVG files can be very large. The PNG is the smallest.

The second is to specify a *figfile*. You probably want to disable the GUI using *interactive=False* in this case. The type of plot file that is made will depend upon the filename suffix. The allowed choices are *.png* (PNG), *.eps* (EPS), and *svg* (SVG).

This latter option is most useful from scripts. For example,

```
default('plotxy')
vis = 'ngc5921.ms'
field = '2'
spw = ''
```

²Also at <http://www.aoc.nrao.edu/~egreisen/scs.ps>

```
xaxis = 'uvdist'
yaxis = 'amp'
interactive=False
figfile = 'ngc5921.uvplot.amp.png'
plotxy()
```

will plot amplitude versus uv-distance in PNG format. No `plotxy` GUI will appear.

ALERT: if you use this option to print to *figfile* with an *iteration* set, you will only get the first plot.

3.3.2.14 Exiting plotxy

You can use the **Quit** button to clear the plot from the window and detach from the MS. You can also dismiss the window by killing it with the X on the frame, which will also detach the MS.

You can also just leave it alone. The plotter pretty much keeps running in the background even when it looks like it's done! You can keep doing stuff in the plotter window, which is where the *overplot* parameter comes in. Note that the *plotcal* task (§ 4.5.1) will use the same window, and can also overplot on the same panel.

If you leave *plotxy* running, beware of (for instance) deleting or writing over the MS without stopping. It may work from a memory version of the MS or crash.

3.3.2.15 Example session using plotxy

The following is an example of interactive plotting and flagging using `plotxy` on the Jupiter 6cm continuum VLA dataset. This is extracted from the script *jupiter6cm.usecase.py* available in the script area.

This assumes that the MS *jupiter6cm.usecase.ms* is on disk with *flagautocorr* already run.

```
default('plotxy')

vis = 'jupiter6cm.usecase.ms'

# The fields we are interested in: 1331+305,JUPITER,0137+331
selectdata = True

# First we do the primary calibrator
field = '1331+305'

# Plot only the RR and LL for now
correlation = 'RR LL'

# Plot amplitude vs. uvdist
xaxis = 'uvdist'
yaxis = 'amp'
```

```
multicolor = 'both'

# The easiest thing is to iterate over antennas
iteration = 'antenna'

plotxy()

# You'll see lots of low points as you step through RR LL RL LR
# A basic clip at 0.75 for RR LL and 0.055 for RL LR will work
# If you want to do this interactively, set
iteration = ''

plotxy()

# You can also use flagdata to do this non-interactively
# (see below)

# Now look at the cross-polar products
correlation = 'RL LR'

plotxy()

#-----
# Now do calibrator 0137+331
field = '0137+331'
correlation = 'RR LL'
xaxis = 'uvdist'
spw = ''
iteration = ''
antenna = ''

plotxy()

# You'll see a bunch of bad data along the bottom near zero amp
# Draw a box around some of it and use Locate
# Looks like much of it is Antenna 9 (ID=8) in spw=1

xaxis = 'time'
spw = '1'
correlation = ''

# Note that the strings like antenna='9' first try to match the
# NAME which we see in listobs was the number '9' for ID=8.
# So be careful here (why naming antennas as numbers is bad).
antenna = '9'

plotxy()

# YES! the last 4 scans are bad. Box 'em and flag.

# Go back and clean up
```

```
xaxis = 'uvdist'
spw = ''
antenna = ''
correlation = 'RR LL'

plotxy()

# Box up the bad low points (basically a clip below 0.52) and flag

# Note that RL,LR are too weak to clip on.

#-----
# Finally, do JUPITER
field = 'JUPITER'
correlation = ''
iteration = ''
xaxis = 'time'

plotxy()

# Here you will see that the final scan at 22:00:00 UT is bad
# Draw a box around it and flag it!

# Now look at what's left
correlation = 'RR LL'
xaxis = 'uvdist'
spw = '1'
antenna = ''
iteration = 'antenna'

plotxy()

# As you step through, you will see that Antenna 9 (ID=8) is often
# bad in this spw. If you box and do Locate (or remember from
# 0137+331) it's probably a bad time.

# The easiest way to kill it:

antenna = '9'
iteration = ''
xaxis = 'time'
correlation = ''

plotxy()

# Draw a box around all points in the last bad scans and flag 'em!

# Now clean up the rest
xaxis = 'uvdist'
correlation = 'RR LL'
antenna = ''
```

```

spw = ''

# You will be drawing many tiny boxes, so remember you can
# use the ESC key to get rid of the most recent box if you
# make a mistake.

plotxy()

# Note that the end result is we've flagged lots of points
# in RR and LL. We will rely upon imager to ignore the
# RL LR for points with RR LL flagged!

```

3.3.3 Plotting antenna positions using plotants

This task is a simple plotting interface (to the `plotxy` functionality) to produce plots of the antenna positions (taken from the *ANTENNA* sub-table of the MS).

The inputs to `plotants` are:

```

# plotants :: Plot the antenna distribution in the local reference frame:
vis      =      '' # Name of input visibility file (MS)
figfile  =      '' # Save the plotted figure to this file

```

3.3.4 Plotting uv-coverages plotuv

A simple way to plot uv-coverages is offered by the task `plotuv`:

```

# plotuv :: Plot the baseline distribution
vis      =      '' # Name of input visibility file (MS)
field    =      '' # Select field using ID(s) or name(s)
antenna  =      '' # Select data based on antenna/baseline
spw      =      '' # Select spectral window/channels
observation =      '' # Select by observation ID(s)
array    =      '' # Select (sub)array(s) by array ID number
maxnpts  =      100000 # Maximum number of points per plot.
colors   =      ['r', 'y', 'g', 'b'] # a list of matplotlib color codes
symb     =      ',' # A matplotlib plot symbol code
ncycles  =      1 # How many times to cycle through colors per
# plot.
figfile  =      '' # Save the plotted figure(s) using this name

```

`plotuv` provides basic selection of data as well as plotting style options. The difference to `plotms` is that `plotuv` is also plotting the Hermitian conjugates of the visibilities which produces the familiar symmetric plots. This is a remedy to the restriction in `plotms` to allow flagging of data. This is achieved via a unambiguous link from a displayed data point to a visibility. Plotting Hermitian conjugates would break this rule in `plotms` and `plotuv` is used instead to plot Hermitian conjugates.

3.4 Data Flagging using flagdata

flagdata can flag measurement sets and calibration tables with an elaborate selection syntax. It also contains auto-flagging routines.

For a full description of flagdata please visit:

<http://www.aoc.nrao.edu/~rurvashi/FlaggerDocs/FlaggerDocs.html>

The inputs to flagdata are:

```
# flagdata :: All-purpose flagging task based on data-selections and flagging modes/algorithms.
vis                =      ''          # Name of MS file or calibration table to flag
mode               =      'manual'    # Flagging mode
  field            =      ''          # Field names or field index
                                     # numbers: '' ==> all, field='0~2,3C286'
  spw              =      ''          # Spectral-window/frequency/channel: '' ==> all, spw='0:17~19'
  antenna          =      ''          # Antenna/baselines: '' ==> all, antenna = '3,VA04'
  timerange        =      ''          # Time range: '' ==> all, timerange='09:14:0~09:54:0'
  correlation      =      ''          # Correlation: '' ==> all, correlation='XX,YY'
  scan             =      ''          # Scan numbers: '' ==> all
  intent           =      ''          # Observation intent: '' ==> all, intent='CAL*POINT*'
  array            =      ''          # (Sub)array numbers: '' ==> all
  uvrange          =      ''          # UV range: '' ==> all;
                                     # uvrange = '0~100klambda', default units=meters
  observation      =      ''          # Observation ID: '' ==> all
  feed             =      ''          # Multi-feed numbers: Not yet implemented
  autocorr         =      False       # Flag auto-correlations

action             =      'apply'     # Action to perform in MS
                                     # and/or in infile (none/apply/calculate)
  display          =      ''          # Display data and/or
                                     # end-of-MS reports at runtime (data/report/both).
  flagbackup       =      True        # Back up the state of flags before the run

savepars           =      False       # Save the current parameters
                                     # to the FLAG_CMD table or to a file
```

vis can take a measurement set or calibration table. Data selection for calibration tables is limited to *field*, *scan*, *time*, *antenna*, *spw*, and *observation*. Since calibration tables do not have a FLAG_CMD table, parameter settings, if requested, can only be saved in external files.

The *mode* parameter (§3.4.2) selects the flagging algorithm and the following are available:

```
list              = list of flagging commands to apply to MS
manual            = flagging based on specific selection parameters
clip              = clip data according to values
quack             = remove/keep specific time range at scan beginning/end
shadow           = remove antenna-shadowed data
```

```

elevation  = remove data below/above given elevations
tfcrop     = automatic identification of outliers on the time-freq plane
rflag      = automatic detection of outliers based on sliding-window RMS filters
extend     = extend and/or grow flags beyond what the basic algorithms detect
summary    = report the amount of flagged data
unflag     = unflag the specified data

```

Flagging will only be applied to the data selection that is performed with the usual selection parameters (§ 2.3). The dataset is iterated-through in chunks (small pieces of data) consisting of one field, one spw, and a user-defined timerange (default is one scan). In addition to the typical antenna, spw, timerange, etc. selections, we would like to point out some addition of the *correlation* syntax for modes *clip*, *tfcrop*, and *rflag*. One can combine correlation products with simple mathematical expressions

```
'ABS', 'ARG', 'RE', 'IM', 'NORM'
```

followed by the polarization products (using an underscore in between “_”)

```
'ALL', 'I', 'XX', 'YY', 'RR', 'LL', 'WVR'
```

‘WVR’ refers to the water vapour radiometer of ALMA data. Note that the operators *ABS*, *ARG*, *RE*, etc. are written only once as the first value. if more than one correlation is given, the operator will be applied to all of them. An example would be

```
correlation='RE_XX,XY'
```

which would select all real XX and XY polarization for flagging.

3.4.1 The action parameter

The keyword *action* controls whether the actual flagging commands will be applied or not and the options are the empty string “”, ‘*apply*’ and ‘*calculate*’.

apply is likely the most popular one as it applies the flags to the MS:

```

action          = 'apply'          # Action to perform in MS and/or in infile
                                   # (none/apply/calculate)
  display       = ''              # Display data and/or end-of-MS reports at runtime
                                   # (data/report/both).
  flagbackup    = True            # Back up the state of flags before the run

```

flagbackup specifies if a backup of the current flags should be saved in the “*.flagversions” file. *display* can be “”, ‘*data*’, ‘*report*’, ‘*both*’ where the empty string “” will report no individual flagging statistics, whereas ‘*data*’ launches an interactive GUI to display data and flags for each chunk to

browse through. The plots are time-frequency planes and both old and new flags are being overlaid for all correlations per baseline. In the GUI, one can step through all chunks for inspection and if the flagging is unsatisfactory, one can exit without applying the flags. If the flagging is acceptable, it is also possible to continue flagging without viewing all chunks (the number of chunks can be very large for typical JVLA and ALMA data sets. *display='report'* lists the flagging statistics at the end of the procedure on the screen and *both* starts the GUI and reports all statistics at the end.

action='calculate' calculates the flags but does not write them to the MS or calibration table. This is useful if one would like to inspect the computed flags in the GUI without a straight application:

```
action          = 'calculate'      # Action to perform in MS and/or in infile
                                     # (none/apply/calculate)
display         = ''              # Display data and/or end-of-MS reports at runtime
                                     # (data/report/both).
```

The empty string *action=""* will do nothing and is useful when the commands themselves shall only be written to the *FLAG_CMD* sub-table or to an external file using the *savepars* parameter to specify the filename.

savepars will save the flagging commands to a file that can be later used for input in *flagdata* via *mode='list'*. It also shares the *flagcmd* syntax and can be used there. The file name is specified by *outfile* and, if empty, the *FLAG_CMD* table in the MS will be populated. A *REASON* can be given by the *reason* keyword which may be useful for bookkeeping as well as for unflagging data that are marked by specific *REASON* keywords.

3.4.2 Flagging Modes

3.4.2.1 Manual Flag/Unflag

```
mode           = 'manual'         # Flagging mode (list/manual/clip/shadow/quack/el
                                     # evation/tfcrop/rflag/extend/unflag/summary)
field          = ''              # Field names or field index numbers: '' ==> all,
                                     # field='0~2,3C286'
spw            = ''              # Spectral-window/frequency/channel: '' ==> all,
                                     # spw='0:17~19'
antenna        = ''              # Antenna/baselines: '' ==> all, antenna
                                     # ='3,VA04'
timerange      = ''              # Time range: '' ==>
                                     # all,timerange='09:14:0~09:54:0'
correlation    = ''              # Correlation: '' ==> all, correlation='XX,YY'
scan           = ''              # Scan numbers: '' ==> all
intent         = ''              # Observation intent: '' ==> all,
                                     # intent='CAL*POINT*'
array          = ''              # (Sub)array numbers: '' ==> all
uvrange        = ''              # UV range: '' ==> all; uvrange ='0~100klambda',
                                     # default units=meters
observation    = ''              # Observation ID: '' ==> all
```

```

feed          =          ''          # Multi-feed numbers: Not yet implemented
autocorr      =          False       # Flag auto-correlations

```

The *'manual'* mode is the most straight-forward of all modes. All visibilities that are selected by the various data selection parameters will be flagged or unflagged, depending on the *action* parameter. *autocorr* is a shorthand for *antenna='*@@@'* to flag all auto correlations in the data.

3.4.2.2 List

```

mode          =          'list'       # Flagging mode (list/manual/clip/shadow/quack/el
#          evation/tfcrop/rflag/extend/unflag/summary)
inppfile      =          ''          # Input ASCII file, list of
#          files or Python list of strings with
#          flag commands.
reason        =          'any'       # Select by REASON types

```

A list of flag commands can be provided through a file or a list of files, specified by the *inppfile* parameter. Each input line may contain a flagging *mode* with data selection parameters as well as parameters that are specific to that *mode*. All parameters that are not set will be reset to their default values (default *mode* is *'manual'*). Each line of this file or list of strings will be taken as a command to the *flagdata* task. This *mode='list'* is similar to the task *flagcmd* with the *inppmode='list'* option.

An example for such a file would be:

```

mode='shadow'
mode='clip' clipminmax=[0,5] correlation='ABS_ALL'
mode='quack' quackmode='end' quackinterval=1.0
antenna='ea01' timerange='00:00:00~01:00:00'
antenna='ea11' timerange='00:00:00~03:00:00' spw='0~4'

```

Alternatively, this can be issued in the task directly like:

```

flagdata(vis='vis',mode='list',
inppfile=["mode='shadow'",
          "mode='clip' clipminmax=[0,5] correlation='ABS_ALL'",
          "mode='quack' quackmode='end' quackinterval=1.0",
          "antenna='ea01' timerange='00:00:00~01:00:00'",
          "antenna='ea11' timerange='00:00:00~03:00:00' spw='0~4'"])

```

or via a variable

```

cmds=["mode='shadow',
      "mode='clip' clipminmax=[0,5] correlation='ABS_ALL'",

```


This option flags shadowed antennas, i.e. when one antenna blocks part of the aperture of a second antenna that is behind the first one. Shadowing can be gradual and the criterion for a shadow flag is when a baseline is shorter than $\text{radius}_1 + \text{radius}_2 - \text{tolerance}$ (where the radii of the antennae are taken from the MS antenna subtable); see Fig. 3.9. *addantenna* may be used to account for shadowing when antennas are not listed in the MS but are physically present. Please read the *flagdata* inline help for the syntax of this option.

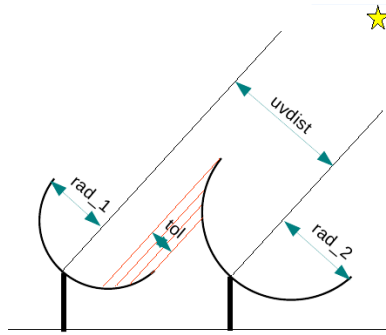


Figure 3.9: This figure shows the geometry used to compute shadowed antennas.

3.4.2.5 Quack

```

mode           = 'quack'           # Flagging mode (list/manual/clip/shadow/quack/
                                # elevation/tfcrop/rflag/extend/unflag/summary)
...
quackinterval = 0.0                # Quack n seconds from scan beginning or end
quackmode     = 'beg'              # Quack mode. 'beg' ==>
                                # beginning of scan. 'endb' ==> end of
                                # scan. 'end' ==> all but
                                # end of scan. 'tail' ==> all but
                                # beginning of scan
quackincrement = False            # Flag incrementally in time?

```

quack is used to remove data at scan boundaries. *quackinterval* specifies the time in seconds to be flagged, and *quackmode* can be *'beg'* to flag the *quackinterval* at the beginning of each selected scan, *'endb'* at the end of scan. *'tail'* flags all but the beginning of scan and *'end'* all but the end of scan. The *quackincrement* is either *True* or *False*, depending if one wishes to flag the *quackinterval* from the first unflagged data in the scan, or from the scan boundaries independent of data being already flagged or not.

3.4.2.6 Elevation

```

mode           = 'elevation'       # Flagging mode (list/manual/clip/shadow/quack/
                                # elevation/tfcrop/rflag/extend/unflag/summary)

```

```

...
lowerlimit    =      0.0      # Lower limiting elevation (in degrees)
upperlimit    =      90.0     # Upper limiting elevation (in degrees)

```

Flagging based on the elevation of the antennae. This may be useful to avoid data taken at very low elevations or close to transit and the *lowerlimit* and *upperlimit* parameters specify the range of good elevations.

3.4.2.7 Tfcrop

```

mode          =  'tfcrop'      # Flagging mode (list/manual/clip/shadow/quack/
                             # elevation/tfcrop/rflag/extend/unflag/summary
                             # )
...
ntime         =  'scan'       # Time-range to use for each chunk (in seconds
                             # or minutes)
combinescans  =  False        # Accumulate data across scans.
datacolumn    =  'DATA'       # Data column on which to operate
                             # (data,corrected,model,residual)
timecutoff    =      4.0      # Flagging thresholds in units of deviation
                             # from the fit
freqcutoff    =      3.0      # Flagging thresholds in units of deviation
                             # from the fit
timefit       =  'line'       # Fitting function for the time direction
                             # (poly/line)
freqfit       =  'poly'       # Fitting function for the frequency direction
                             # (poly/line)
maxnpieces    =      7        # Number of pieces in the polynomial-fits (for
                             # 'freqfit' or 'timefit' = 'poly')
flagdimension = 'freqtime'    # Dimensions along which to calculate fits
                             # (freq/time/freqtime/timefreq)
usewindowstats = 'none'      # Calculate additional flags using sliding
                             # window statistics (none,sum,std,both)
halfwin       =      1        # Half-width of sliding window to use with
                             # 'usewindowstats' (1,2,3).
extendflags   =      True     # Extend flags along time,
                             # frequency and correlation.

```

TFCrop is an autoflag algorithm that detects outliers on the 2D time-frequency plane, and can operate on un-calibrated data (non bandpass-corrected). The original implementation of this algorithm is described in NCRA Technical Report 202 (Oct 2003).

The algorithm iterates through the data in chunks of time. For each chunk, the result of user-specified visibility-expressions are organized as 2D time-frequency planes, one for each baseline and correlation-expression result, and the following steps are performed.

1. Calculate a bandshape template : Average the data across time, to construct an average bandpass. Construct an estimate of a clean bandpass (without RFI) via a robust piece-wise polynomial fit to the average bandpass shape.

Note : A robust fit is computed in up to 5 iterations. It begins with a straight line fit across the full range, and gradually increases to 'maxnpieces' number of pieces with third-order polynomials in each piece. At each iteration, the stddev between the data and the fit is computed, values beyond N-stddev are flagged, and the fit and stddev are re-calculated with the remaining points. This stddev calculation is adaptive, and converges to a value that reflects only the data and no RFI. At each iteration, the same relative threshold is applied to detect flags, and this results in a varying set of flagging thresholds, that allows deep flagging only when the fit represents the true data best. Iterations stop when the stddev changes by less than 10%, or when 5 iterations are completed.

The resulting clean bandpass is a fit across the base of RFI spikes.

2. Divide out this clean bandpass function from all timesteps in the current chunk. Now, any data points that deviate from a mean of 1 can be considered RFI. This step helps to separate narrow-band RFI spikes from a smooth but varying bandpass, in situations where a simple range-based clipping will flag good sections of the bandpass.
3. Perform iterative flagging (robust flagging) of points deviating from a value of 1.

Flagging is done in up to 5 iterations. In each iteration, for every timestep, calculate the stddev of the bandpass-flattened data, flag all points further than N times stddev from the fit, and recalculate the stddev. At each iteration, the same relative threshold is applied to detect flags. Optionally, use sliding-window based statistics to calculate additional flags.

4. Repeat steps 1 and 3, but in the other direction (i.e. average the data across frequency, calculate a piece-wise polynomial fit to the average time-series, and find flags based on deviations w.r.to this fit.)

The default parameters of the `tfcrop` implementation are optimized for strong narrow-band RFI (see, e.g. Fig. 3.10). With broad-band RFI, the piece-wise polynomial can sometimes model it as part of the band-shape, and therefore not detect it as RFI. In this case, reducing the maximum number of pieces in the polynomial can help. This algorithm usually has trouble with noisy RFI that is also extended in time or frequency, and additional statistics-based flagging is recommended (via the 'usewindowstats' parameter). It is often required to set up parameters separately for each spectral-window.

If frequency ranges of known astronomical spectral lines are known *a-priori*, they can be protected from automatic flagging by de-selecting those frequency-ranges via the 'spw' data-selection parameter.

The `extendflag` parameter will clean up small portions of data between flagged data points along time and/or frequency when more than 50% of all timeranges or 80% of all channels are already flagged. It will also extend the flags to the other polarizations. Alternatively, `mode='extend'` can be used (Fig. 3.12).

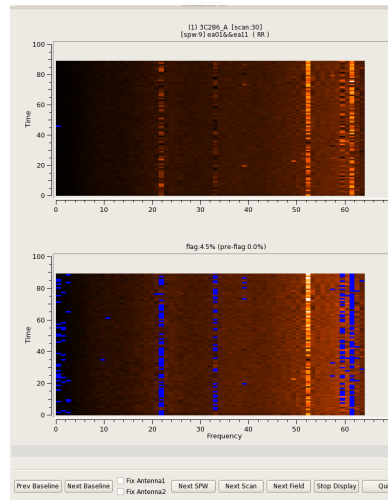


Figure 3.10: This screenshot represents a run where 'tfcrop' was run on a spw='9' with mainly narrow-band RFI. RIGHT : An example of protecting a spectral line (in this case, demonstrated on an RFI spike) by setting the spw-selection to spw='0:0 45:53 63'. In both figures, the top row indicates the data before flagging, and the bottom row after flagging.

3.4.2.8 Rflag

```

mode          = 'rflag'      # Flagging mode (list/manual/clip/shadow/quack/
                        # elevation/tfcrop/rflag/extend/unflag/summary
                        # )
...
ntime         = 'scan'      # Time-range to use for each chunk (in seconds
                        # or minutes)
combinescans  = False       # Accumulate data across scans.
datacolumn   = 'DATA'      # Data column on which to operate
                        # (data,corrected,model,residual)
winsize      = 3           # Number of timesteps in the sliding time
                        # window [aips:fparm(1)]
timedev      = ''          # Time-series noise estimate [aips:noise]
freqdev      = ''          # Spectral noise estimate [aips:scutoff]
timedevscale = 5.0         # Threshold scaling for timedev [aips:fparm(9)]
freqdevscale = 5.0         # Threshold scaling for freqdev
                        # [aips:fparm(10)]
spectralmax  = 1000000.0   # Flag whole spectrum if freqdev is greater
                        # than spectralmax [aips:fparm(6)]
spectralmin  = 0.0         # Flag whole spectrum if freqdev is less than
                        # spectralmin [aips:fparm(5)]
extendflags  = True        # Extend flags along time, frequency and correlation.

```

RFlag is an autoflag algorithm based on a sliding window statistical filter. The RFlag algorithm was originally developed by Eric Greisen in AIPS (31DEC11). AIPS documentation : Subsection E.5 of the AIPS cookbook (Appendix E : Special Considerations for JVLA data calibration and imaging in AIPS³)

In RFlag, the data is iterated-through in chunks of time, statistics are accumulated across time-chunks, thresholds are calculated at the end, and applied during a second pass through the dataset.

The CASA implementation also optionally allows a single-pass operation where statistics and thresholds are computed and also used for flagging, within each time-chunk (defined by 'ntime' and 'combinescans').

For each chunk, calculate local statistics, and apply flags based on user supplied (or auto-calculated) thresholds.

1. Time analysis (for each channel)
 - (a) Calculate local rms of real and imag visibilities, within a sliding time window
 - (b) Calculate the median rms across time windows, deviations of local rms from this median, and the median deviation
 - (c) Flag if local rms is larger than $\text{timedevscale} \times (\text{medianRMS} + \text{medianDev})$
2. Spectral analysis (for each time)
 - (a) Calculate avg of real and imag visibilities and their rms across channels
 - (b) Calculate the deviation of each channel from this avg, and the median-deviation
 - (c) Flag if deviation is larger than $\text{freqdevscale} \times \text{medianDev}$

The *extendflag* parameter will clean up small portions of data between flagged data points along time and/or frequency when more than 50% of all timeranges or 80% of all channels are already flagged. It will also extend the flags to the other polarizations. Alternatively, *mode='extend'* can be used.

Some examples (also see Fig. 3.11):

1. Calculate thresholds automatically per scan, and use them to find flags. Specify scale-factor for time-analysis thresholds, use default for frequency.

```
flagdata('my.ms', mode='rflag', spw='9', timedevscale=4.0, writeflags=True)
```

2. Supply noise-estimates to be used with default scale-factors.

```
flagdata(vis='my.ms', mode='rflag', spw='9', timedev=0.1, freqdev=0.5, writeflags=True);
```

³<http://www.aips.nrao.edu/cook.html#CEE>

Two-passes. This replicates the usage pattern in AIPS.

- The first pass saves commands in an output text files, with auto-calculated thresholds. Thresholds are returned from `rflag` only when `writeflags=False` (calc-only mode). The user can edit this file before doing the second pass, but the python-dictionary structure must be preserved.
- The second pass applies these commands (`writeflags=True`).

```
flagdata(vis='my.ms', mode='rflag', spw='9,10',
flagdata(vis='my.ms', mode='rflag', spw='9,10',
        timedev='tdevfile.txt', freqdev='fdevfile.txt', writeflags=True);
```

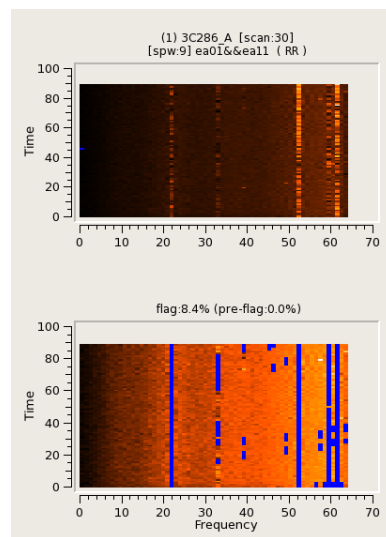


Figure 3.11: Example of *rflag* on narrow-band RFI

3.4.2.9 Extend

```
mode           = 'extend'      # Flagging mode (list/manual/clip/shadow/quack/evaluation/tfcrop/rflag/extend/unflag/summary)
field          = ''           # Field names or field index numbers: '' ==> all,
                             # field='0~2,3C286'
spw            = ''           # Spectral-window/frequency/channel: '' ==> all,
                             # spw='0:17~19'
antenna        = ''           # Antenna/baselines: '' ==> all, antenna
                             # ='3,VA04'
timerange      = ''           # Time range: '' ==>
                             # all,timerange='09:14:0~09:54:0'
correlation    = ''           # Correlation: '' ==> all, correlation='XX,YY'
scan           = ''           # Scan numbers: '' ==> all
```

```

intent      =      ''      # Observation intent: '' ==> all,
                        # intent='CAL*POINT*'
array       =      ''      # (Sub)array numbers: '' ==> all
uvrange     =      ''      # UV range: '' ==> all; uvrange ='0~100klambda',
                        # default units=meters
observation =      ''      # Observation ID: '' ==> all
feed        =      ''      # Multi-feed numbers: Not yet implemented
ntime       =      'scan'  # Time-range to use for each chunk (in seconds or
                        # minutes)
combinescans =      False  # Accumulate data across scans.
extendpols  =      True   # If any correlation is flagged, flag all
                        # correlations
growtime    =      50.0   # Flag all 'ntime' integrations if more than X%
                        # of the timerange is flagged (0-100)
growfreq    =      50.0   # Flag all selected channels if more than X% of
                        # the frequency range is flagged(0-100)
growaround  =      False  # Flag data based on surrounding flags
flagneartime =      False # Flag one timestep before and after a flagged
                        # one (True/False)
flagnearfreq =      False # Flag one channel before and after a flagged one
                        # (True/False)

```

Although the modes *tfcrop* and *rflag* already have *extendflag* parameters, some autoflagging algorithms may still leave small islands of unflagged data behind, data that are surrounded by flagged visibilities in the time-frequency space. Although the algorithm may deem these visibilities as good ones, they are frequently affected by low-level RFI that spills from the adjacent, flagged points and one may wish to clean those up.

ntime specifies the time ranges over which to clean up, e.g. *'1.5min'* or *'scan'* which checks on all data within a scan. To span time ranges larger than scans, one can set *combinescans* to *True*.

extendpols=True would extend all flags to all polarization products when at least one of them is flagged.

growtime flags the entire time range for a flagged channel, when a certain fraction of flagged time intervals is exceeded.

growfreq is similar but extends the flags in frequency when a given fraction of channels is already flagged.

growaround checks for flagged data points in the time-frequency domain that neighbor a datum. The threshold is four data points. If more surrounding points are flagged, the central datum will be flagged, too.

flagneartime flags adjacent data points along the time axis, around a flagged datum

flagnearfreq flags neighboring channels.

For an example, see Fig. 3.12.

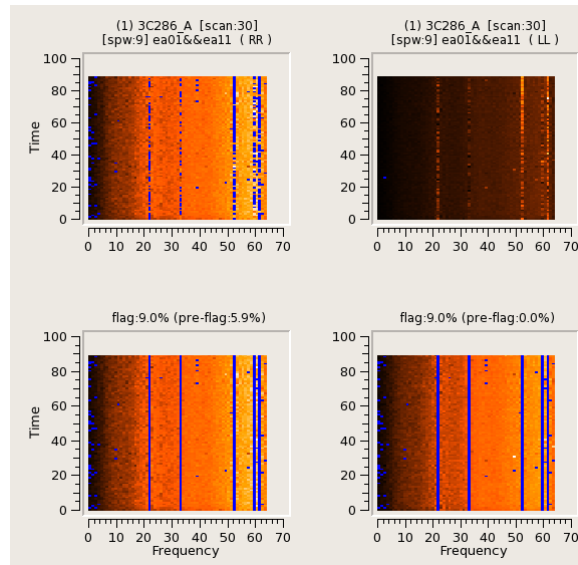


Figure 3.12: This screenshot represents a run where 'tfcrop' was run only on 'ABS_RR' (top row) and followed by an extension along time and correlations (bottom row).

3.4.2.10 Unflag

```

mode                = 'unflag'          # Flagging mode (list/manual/clip/shadow/quack/
                                # elevation/tfcrop/rflag/extend/unflag/summary
                                # )
field               = ''                # Field names or field index numbers: ''=>all,
                                # field='0~2,3C286'
spw                 = ''                # spectral-window/frequency/channel
antenna             = 'ea01'           # antenna/baselines: ''=>all, antenna
                                # ='3,VA04'
timerange           = ''                # time range:
                                # ''=>all,timerange='09:14:0~09:54:0'
correlation         = ''                # Select data based on correlation
scan                = ''                # scan numbers: ''=>all
intent              = ''                # Select data based on observation intent:
                                # ''=>all
feed                = ''                # multi-feed numbers: Not yet implemented
array               = ''                # (sub)array numbers: ''=>all
uvrange             = ''                # uv range: ''=>all; uvrange ='0~100klambda',
                                # default units=meters
observation          = ''                # Select data based on observation ID: ''=>all

```

The selection data will be unflagged.

3.4.2.11 Summary

```

mode          = 'summary'      # Flagging mode (list/manual/clip/shadow/quack/
                                # elevation/tfcrop/rflag/extend/unflag/summary
                                # )
...
minrel        = 0.0            # minimum number of flags (relative)
maxrel        = 1.0            # maximum number of flags (relative)
minabs        = 0              # minimum number of flags (absolute)
maxabs        = -1            # maximum number of flags (absolute). Use a
                                # negative value to indicate infinity.
spwchan       = False         # Print summary of channels per spw
spwcorr       = False         # Print summary of correlation per spw
basecnt       = False         # Print summary counts per baseline

```

This mode reports the number of rows and data points that are flagged. The selection of reported points can be restricted (see inline help for details).

`mode='summary'` can also report back a dictionary if the task is run as

```
s = flagdata(..., mode='summary')
```

with a variable assigned, here 's'.

3.5 Command-based flagging using flagcmd

The task `flagcmd` will flag the visibility data set or calibration table based on a specified set of flagging commands using a flagging syntax (see § 3.5.3). These commands can be input from the `FLAG_CMD` MS table, from a `Flag.xml` SDM table, from an ascii file, or from input python strings. Facilities for manipulation, listing, or plotting of these flags are also provided.

The inputs to `flagcmd` are:

```

# flagcmd :: Flagging task based on batches of flag-commands
vis          = ''             # Name of MS file or calibration table to flag
inpmode      = 'table'       # Input mode for flag commands(table/list/xml)
  inpfile    = ''           # Source of flag commands
  tablerows  = []           # Rows of inpfile to read
  reason     = 'any'        # Select by REASON types
  useapplied = False        # Select commands whose rows
                                # have APPLIED column set to True

action       = 'apply'       # Action to perform in MS and/or in inpfile
                                # (apply/unapply/list/plot/clear/extract)
  flagbackup = True         # Automatically backup the
                                # FLAG column before execution

savepars     = False        # Save flag commands to the MS or to a file

```

The default input mode is `inpmode='table'` which directs the task to input flag commands from the `FLAG_CMD` internal MS table. See § 3.5.1 for more options.

The default operation mode is `action='apply'` directing the task to apply relevant flagging commands to the MS data main table. See § 3.5.2 for more options.

See § 3.5.3 for a description of the flagging command syntax.

It is possible to flag calibration tables using `flagcmd`, although we recommend using the `flagdata` task for this.

When using `flagcmd` to flag calibration tables, only the `apply` and `list` actions are supported. Because calibration tables do not have a `FLAG_CMD` sub-table, the default `inpmode='table'` can only be used if an MS is given in the `inpfile` parameter so that flags from the MS are applied to the calibration table directly. Otherwise, the flag commands must be given using `inpmode='list'`, either from a file or from a list of strings.

3.5.1 Input modes `inpmode`

The `inpmode` parameter selects options for the input mode for the flagging commands.

Available `inpmode` options are:

- `'table'` — input from MS table (§ 3.5.1.1)
- `'list'` — input from ASCII file or from a list of strings (§ 3.5.1.2)
- `'xml'` — input from XML table (§ 3.5.1.3)

3.5.1.1 Input mode `'table'`

The default input mode is `inpmode='table'` which directs the task to input flag commands from a `FLAG_CMD` MS table. This has the sub-parameters:

```

inpmode          = 'table'          # Input mode for flag commands(table/list/xml)
  inpfile        = ''                # Source of flag commands
  tablerows      = []                # Rows of inpfile to read
  reason        = 'any'              # Select by REASON types
  useapplied     = False             # Select commands whose rows
                                      # have APPLIED column set to
                                      # True

```

If `inpfile = ''` then it will look for the `FLAG_CMD` table in the MS given by `vis`. You can use this sub-parameter to direct the task to look directly at another table.

The `tablerows` sub-parameter is a simple Python list of the row numbers of the table to consider in processing flags. The default is all rows.

The *useapplied* sub-parameter toggles whether only flag commands marked as not having been applied are considered (the default), or to allow (re)processing using all commands.

The *reason* sub-parameter selects the *REASON* type to process. The default *'any'* means all commands, note that *reason=""* would only select flags who have a blank *REASON* column entry.

One use case is to read the flag commands from the *FLAG_CMD* of an MS and apply them to a calibration table given in the parameter *vis*. Example:

```
flagcmd(vis='cal-X54.B1', inpmode='table',
        inpfile='uid__A002_X2a5c2f_X54.ms', action='apply')
```

3.5.1.2 Input flag mode *'list'*

This mode allows one to insert a list of strings with flagging commands, the name of a file or a list of filenames that contains these commands equivalent to the *mode='list'* in *flagdata* (§3.4.2.2). E.g. a file *flags.txt* that contains

```
scan='1~3' mode='manual'
mode='clip' clipminmax=[0,2] correlation='ABS_XX' clipoutside=False
spw='9' mode='tfcrop' correlation='ABS_YY' ntime=51.0
mode='extend' extendpols=True
```

can be called via

```
flagcmd(vis,inpmode='list',inpfile='flags.txt')
```

Alternatively, the individual flagging commands can be directly provided in the call itself like

```
inpfile=["scan='1~3' mode='manual'",
        "mode='clip' clipminmax=[0,2] correlation='ABS_XX' clipoutside=False",
        "spw='9' mode='tfcrop' correlation='ABS_YY' ntime=51.0",
        "mode='extend' extendpols=True"]
```

3.5.1.3 Input flag mode *'xml'*

The input mode *inpmode='xml'* directs the task to input flag commands from a XML SDM online flagging *Flag.xml* file. When set this opens the sub-parameters:

```
inpmode      =      'xml'      # Input mode for flag commands(table/list/xml)
tbuffer      =      0.0        # Time buffer (sec) to pad flags
ants         =      ''         # Allowed flag antenna names to select by
reason       =      'any'      # Select by REASON types
```


The *flagbackup* toggle sets whether a new copy of the MS main table *FLAG* column is written to the *.flagversions* backup directory for that MS before the requested flagging operation.

3.5.2.2 Unapply flags — *action* option *'unapply'*

The *unapply* option allows unflagging of data based on the selected flag commands. This choice opens the sub-parameters:

```

action          = 'unapply'      # Action to perform in MS and/or in infile
                                   # (apply/unapply/list/plot/clear/extract)
    flagbackup  =      True      # Automatically backup the
                                   # FLAG column before execution

```

As in *action='apply'*, it is possible to make a backup to the **.flagversions* file by using *flagbackup=True*.

In order to guarantee that only the data selected in the command is unapplied, the framework will first unapply the selected rows and then re-apply the overlapping data that got unapplied in the first pass. This is a true unapply action, but it will take longer to process because it will re-apply all the remaining commands that have APPLIED = True!

3.5.2.3 List flags — *action='list'*

The *'list'* option will give a listing of the flagging commands. This choice opens the sub-parameters:

```

action          =      'list'    # Action to perform in MS and/or in infile
                                   # (apply/unapply/list/plot/clear/extract)
savepars        =      True      # Save flag commands to the MS or to a file
    outfile      =      ''       # Name of output file to save commands

```

This action lists the commands on the screen without applying them. One can save the flagging script to an file specified in the *outfile* parameter when *savepars=True*. If *outfile* is empty, it will save the commands to the MS given in *vis*.

The format of the listing output depends on the source of the flagging commands. A set of flagging commands specified through *inpmode='list'* will be listed directly. The flagging commands extracted through *inpmode='table'* will reflect the columns in the table:

```
'Row', 'Timerange', 'Reason', 'Type', 'Applied', 'Lev', 'Sev', 'Command'
```

while commands from *inpmode='xml'* will be shown with the SDM XML table fields:

```
'Key', 'FlagID', 'Antenna', 'Reason', 'Timerange'
```

3.5.2.4 Plot flags — *action='plot'*

The *'plot'* option will produce a graphical plot of flags of time versus antenna. This choice opens the sub-parameters:

```
action          =      'plot'          # Action to perform in MS and/or in infile
                                     # (apply/unapply/list/plot/clear/extract)
      plotfile   =      ''            # Name of output file to save plot
```

This is only useful for online flags or general flag commands that are specified by antenna plus timerange using the standard *REASON* codes that are known SDM *Flag.xml* enumerations.

If the *plotfile* sub-parameter is non-blank, then a plotfile will be made with that name instead of appearing in a matplotlib plotter window on the users workstation.

ALERT: The plotted enumerations are currently only those known to be allowed JVLA online flags as of 15 April 2011, and include:

```
'FOCUS', 'SUBREFLECTOR', 'OFF SOURCE', 'NOT IN SUBARRAY'
```

with all others being plotted as *'Other'*.

3.5.2.5 Clear flags — *action='clear'*

The *'clear'* action will delete selected rows from the *FLAG_CMD* MS table. This choice opens the sub-parameters:

```
action          =      'clear'        # Action to perform in MS and/or in infile
                                     # (apply/unapply/list/plot/clear/extract)
      clearall   =      False         # Delete all rows from FLAG_CMD
      rowlist    =      []            # FLAG_CMD rows to clear
```

The *rowlist* sub-parameter is a simple Python list of the row numbers of the table to consider in processing flags. The default is a blank list which indicates the desire to clear all rows.

In either case, if *clearall=False* then nothing will happen by default as a safeguard. If *clearall=True*, then a blank list will direct the deletion of the selected rows from the table.

ALERT: Use this option with care. You can easily mess up the *FLAG_CMD* table.

3.5.2.6 Extract Flag Commands — *action='extract'*

The *'extract'* option will return the internal flagging dictionary to python:

```
action          =      'extract'      # Action to perform in MS and/or in infile
                                     # (apply/unapply/list/plot/clear/extract)
```

The value can be returned to a variable like:

```
myflagd = flagcmd(vis=msfile,useapplied=True,action='extract')
```

3.5.3 Flagging command syntax

A flagging command syntax has been devised to populate the *COMMAND* column of the *FLAG_CMD* table and to direct the operation of the `flagcmd` task.

The syntax is similar to `flagdata`, so please check *help flagdata* for more info.

You can also use *help flagcmd* inside casapy for this syntax guide also.

Commands are a string (which may contain internal "strings") consisting of *KEY=VALUE* pairs separated by whitespace (see examples below).

NOTE: There should be no whitespace between KEY=VALUE or within each KEY or VALUE, since the simple parser first breaks command lines on whitespace, then on "=".

Each key should only appear once on a given command line/string

There is an implicit "mode" for each command, with the default being 'manual' if not given.

Comment lines can start with '#' and will be ignored.

1. Data selection parameters (used by all flagging modes, see also §2.3)

```
timerange=''
antenna=''
spw=''
correlation=''
field=''
scan=''
feed=''
array=''
uvrange=''
intent=''
observation=''
```

Note: a command consisting only of selection key-value pairs is a basic "manual" operation, i.e. flag the data meeting the selection

2. Modes specific parameters with default values (for further details, refer to the task `flagdata`, §3.4.2).

- (a) Mode *manual*

```
autocorr=False
```

- (b) Mode *clip*

- (c) Mode *manual*

```
datacolumn='DATA'
clipminmax=[]
clipoutside=True
channelavg=False
```

```
clipzeros=False
timeavg=False
timebin=''
```

(d) Mode *shadow*

```
tolerance=0.0
addantenna=''
```

(e) Mode *quack*

```
quackinterval=0.0
quackmode='beg'
quackincrement=False
```

(f) Mode *elevation*

```
lowerlimit=0.0
upperlimit=90.0
```

(g) Mode *tfcrop*

```
ntime='scan'
combinescans=False
datacolumn='DATA'
timecutoff=4.0
freqcutoff=3.0
timefit='line'
freqfit='poly'
maxnpieces=7
flagdimension='freqtime'
usewindowstats='none'
halfwin=1
```

(h) Mode *extend*

```
ntime='scan'
combinescans=False
extendpols=True
growtime=50.0
growfreq=50.0
growaround=False
flagneartime=False
flagnearfreq=False
```

(i) Mode *rflag*

```
ntime='scan'
combinescans=False
datacolumn='DATA'
winsize=3
timedev=''
freqdev=''
timedevscale=5.0
freqdevscale=5.0
spectralmax=1000000.0
spectralmin=0.0
```

(j) Mode *unflag*

3. Basic elaboration options for online and interface use

```
id=''          # flag ID tag (not necessary)
reason=''     # reason string for flag
flagtime=''  # a timestamp for when this flag was generated (for
              user history use)
```

NOTE: there is no *flagtime* column in *FLAG_CMD* at this time, but we will propose to add this as an optional column

NOTE: These are currently ignored and not used

4. Extended elaboration options for online and interface use Note: these are *FLAG_CMD* columns, but their use is not clear but included here for compatibility and future expansion

```
level=N       # flagging "level" for flags with same reason
severity=N    # Severity code for the flag, on a scale of 0-10 in order
              of increasing severity; user specified
```

3.6 Browse the Data

The `browsetable` task is available for viewing data directly (and handles all CASA tables, including Measurement Sets, calibration tables, and images). This task brings up the CASA Qt `casabrowser`, which is a separate program. You can launch this from outside `casapy`.

The default inputs are:

```
# browsetable :: Browse a table (MS, calibration table, image)
tablename      =      '' # Name of input table
```

Currently, its single input is the *tablename*, so an example would be:

```
browsetable('ngc5921.ms')
```

For an MS such as this, it will come up with a browser of the *MAIN* table (see Fig 3.13). If you want to look at sub-tables, use the tab **table keywords** along the left side to bring up a panel with the sub-tables listed (Fig 3.14), then choose (left-click) a table and **View:Details** to bring it up (Fig 3.15). You can left-click on a cell in a table to view the contents.

Note that one useful feature is that you can Edit the table and its contents. Use the *Edit table* choice from the **Edit** menu, or click on the **Edit** button. Be careful with this, and make a backup copy of the table before editing!

The screenshot shows the 'Table Browser' window for a file named 'ngc5921.ms'. The main area displays a table with the following data:

	UVW	FLAG	LAG_CATEGOR	WEIGHT	SIGMA	ANTENNA1	ANTENNA2	ARRAY_ID	DATA
0	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	1	1	0	0
1	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	27	27	0	0
2	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	7	7	0	0
3	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	2	2	0	0
4	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	11	11	0	0
5	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	17	17	0	0
6	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	9	9	0	0
7	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	19	19	0	0
8	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	20	20	0	0
9	[0, 0, 0]	[2, 63] Boolean	[0, 0, 0] Boolean	[23814, 23814]	[0.0514344, 0...	18	18	0	0

The interface also includes a 'PAGE NAVIGATION' section at the bottom with buttons for 'First', '<< [1 / 23] >>', and 'Last', along with a 'Go' button and a 'Loading 1000 rows.' indicator.

Figure 3.13: *browsetable*: The browser displays the main table within a frame. You can scroll through the data (x =columns of the *MAIN* table, and y =the rows) or select a specific page or row as desired. By default, 1000 rows of the table are loaded at a time, but you can step through the MS in batches.

Use the *Close Tables and Exit* option from the **Files** menu to quit the **casabrowser**.

There are a lot of features in the **casabrowser** that are not fully documented here. Feel free to explore the capabilities such as plotting and sorting!

ALERT: You are likely to find that the **casabrowser** needs to get a table lock before proceeding. Use the **clearstat** command to clear the lock status in this case.

The screenshot shows the 'Table Browser' window for the file 'ngc5921.ms'. The window has a menu bar (File, Edit, View, Tools, Export, Help) and a toolbar with various icons. The main area displays a table with the following columns: Keyword, Type, Value, and Extra Information. The table is organized into three sections: 'table data', 'table keywords', and 'field keywords'. The 'table keywords' section is currently selected, showing a list of 14 keywords and their corresponding table paths and row counts.

	Keyword	Type	Value	Extra Information
1	MS_VERSION	Float	2	
2	ANTENNA	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/ANTENNA	Subtable has 28 rows.
3	DATA_DESCRIPTION	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/DATA_D...	Subtable has 1 rows.
4	FEED	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/FEED	Subtable has 28 rows.
5	FLAG_CMD	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/FLAG_C...	Subtable has no rows.
6	FIELD	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/FIELD	Subtable has 3 rows.
7	HISTORY	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/HISTORY	Subtable has 379 rows.
8	OBSERVATION	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/OBSER...	Subtable has 1 rows.
9	POINTING	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/POINTING	Subtable has no rows.
10	POLARIZATION	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/POLARI...	Subtable has 1 rows.
11	PROCESSOR	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/PROCE...	Subtable has no rows.
12	SPECTRAL_WINDOW	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/SPECTR...	Subtable has 1 rows.
13	STATE	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/STATE	Subtable has no rows.
14	SOURCE	Table	/home/sandrock/smyers/Testing/2008-03/ngc5921_regression/ngc5921.ms/SOURCE	Subtable has 3 rows.

Figure 3.14: *browstable*: You can use the tab for *Table Keywords* to look at other tables within an MS. You can then double-click on a table to view its contents.

The screenshot shows a window titled "Table Browser" with a menu bar (File, Edit, View, Tools, Export, Help) and a toolbar. The main area displays a table with the following data:

	DIRECTION	ROPER_MOTIOI	LIBRATION_GRC	CODE	INTERVAL	NAME	NUM_LINES	SOURCE_ID	CTRLAL_W
0	[-2.74393, 0.5...	[0, 0]	-1		1.7976931348...	1331+305000...	1	0	0
1	[-2.42045, 0.1...	[0, 0]	-1		1.7976931348...	1445+099000...	1	1	0
2	[-2.2602, 0.08...	[0, 0]	-1		1.7976931348...	N5921_2	1	2	0

At the bottom of the window, there is a "PAGE NAVIGATION" section with buttons for "First", "<< [1/1] >>", and "Last", along with a "Go" button and a "Loading 1000 rows." indicator.

Figure 3.15: browsetable: Viewing the *SOURCE* table of the MS.

Chapter 4

Synthesis Calibration

This chapter explains how to calibrate interferometer data within the CASA task system. Calibration is the process of determining the net complex correction factors that must be applied to each visibility in order to make them as close as possible to what an idealized interferometer would measure,

such that when the data is imaged an accurate picture of the sky is obtained. This is not an arbitrary process, and there is a philosophy behind the CASA calibration methodology (see § 4.2.1 for more on this). For the most part, calibration in CASA using the tasks is not too different than calibration in other packages such as AIPS or Miriad, so the user should not be alarmed by cosmetic differences such as task and parameter names!

Inside the Toolkit:

The workhorse for synthesis calibration is the `cb` tool.

4.1 Calibration Tasks

Alert: The calibration table format changed in CASA 3.4. CASA 4.2 is the last version that will support the `caltabconvert` function that provides conversions from the pre-3.4 caltable format to the modern format; it will be removed for CASA 4.3. In general, it is best to recalculate calibration using CASA 3.4 or later.

Alert: In CASA 4.2 the `gaincurve` and `opacity` parameters have been removed from all calibration tasks (as advertised in 4.1). These calibration types are supported via the `gencal` task.

Alert: As part of continuing development of a more flexible and improved interface for specifying calibration for `apply`, a new parameter has been introduced in `applycal` and the solving tasks: `docallib`. This parameter toggles between use of the traditional calibration `apply` parameters (`gaintable`, `gainfield`, `interp`, `spwmap`, and `calwt`), and a new `callib` parameter which currently provides access to the *experimental* Cal Library mechanism, wherein calibration instructions are stored in a file. The default remains `docallib=False` in CASA 4.5, and this reveals the traditional `apply` parameters which continue to work as always, and the remainder of this chapter is still written using `docallib=F`. Users interested in the Cal Library mechanism's flexibility are encouraged to try it and report any problems; see Appendix G for information on how to use it, including how to

convert traditional `applycal` to Cal Library format. Note also that `plotms` and `mstransform` now support use of the Cal Library to enable on-the-fly calibration when plotting and generating new MSs.

The standard set of `calibration` solving tasks (to produce calibration tables) are:

- `bandpass` — complex bandpass (B) calibration solving, including options for channel-binned or polynomial solutions (§ 4.4.2),
- `gaincal` — complex gain (G,T) and delay (K) calibration solving, including options for time-binned or spline solutions. (§ 4.4.3),
- `polcal` — polarization calibration including leakage and angle (§ 4.4.5),
- `blcal` — *baseline-based* complex gain or bandpass calibration (§ 4.4.6).

There are helper tasks to create, manipulate, and explore calibration tables:

- `accum` — Accumulate incremental calibration solutions into a cumulative cal table (§ 4.5.5) (**ALERT:** The `accum` task is generally no longer recommended for most calibration scenarios. Please write to the NRAO CASA helpdesk if you need support using `accum`.),
- `applycal` — Apply calculated calibration solutions (§ 4.6.1),
- `clearcal` — Re-initialize the calibration for a visibility dataset (§ 4.6.3),
- `fluxscale` — Bootstrap the flux density scale from standard calibration sources (§ 4.4.4),
- `listcal` — List calibration solutions (§ 4.5.2),
- `plotcal` — Plot calibration solutions (§ 4.5.1),
- `setjy` — Compute model visibilities with the correct flux density for a specified source (§ 4.3.5),
- `smoothcal` — Smooth calibration solutions derived from one or more sources (§ 4.5.4),
- `split` — Write out new MS containing calibrated data from a subset of the original MS (§ 4.7.1).

There are some development versions of calibration and utility tasks that are recently added to the suite:

- `calstat` — Statistics of calibration solutions (§ 4.5.3),
- `cvel` — Regrid a spectral MS onto a new frequency channel system (§ 4.7.7),
- `gencal` — Create a calibration tables from metadata such as antenna position offsets, gain-curves and opacities (§ 4.3.6),

- `wvrgcal` — Generate a gain table based on Water Vapour Radiometer data (for ALMA use - § 4.3.8),
- `hanningsmooth` — Apply a Hanning smoothing filter to spectral-line uv data (§ 4.7.3),
- `mstransform` — *experimental* Task to combine `cvel`, `hanningsmooth`, `split` operations in a single step (§ 4.7.4),
- `uvcontsub` — Carry out uv-plane continuum fitting and subtraction (§ 4.7.6),
- `uvmodelfit` — Fit a component source model to the uv data (§ 4.7.8),
- `uvsub` — Subtract the transform of a model image from the uv data (§ 4.7.5),
- `statwt` — Recalculate the data weights based on their scatter (§ 4.7.9),
- `conjugatevis` — Change the signs of visibility phases (§ 4.7.10).

These are not yet full-featured, and may have only rudimentary controls and options.

The following sections outline the use of these tasks in standard calibration processes.

Information on other useful tasks and parameter setting can be found in:

- `listobs` — summary of a MS (§ 2.2.7),
- `listvis` — list data in a MS (§ 2.2.9),
- `plotms` — prototype next-generation X-Y plotting and editing (§ 3.3.1),
- `plotxy` — previous generation X-Y plotting and editing (§ 3.3.2),
- `plotweather` — plot the weather information of an MS and calculate atmospheric opacities (§ 4.3.4.1),
- `flagdata` — non-interactive data flagging (§ 3.4),
- data selection — general data selection syntax (§ 2.3).

4.2 The Calibration Process — Outline and Philosophy

A work-flow diagram for CASA calibration of interferometry data is shown in Figure 4.1. This should help you chart your course through the complex set of calibration steps. In the following sections, we will detail the steps themselves and explain how to run the necessary tasks and tools.

This can be broken down into a number of discrete phases:

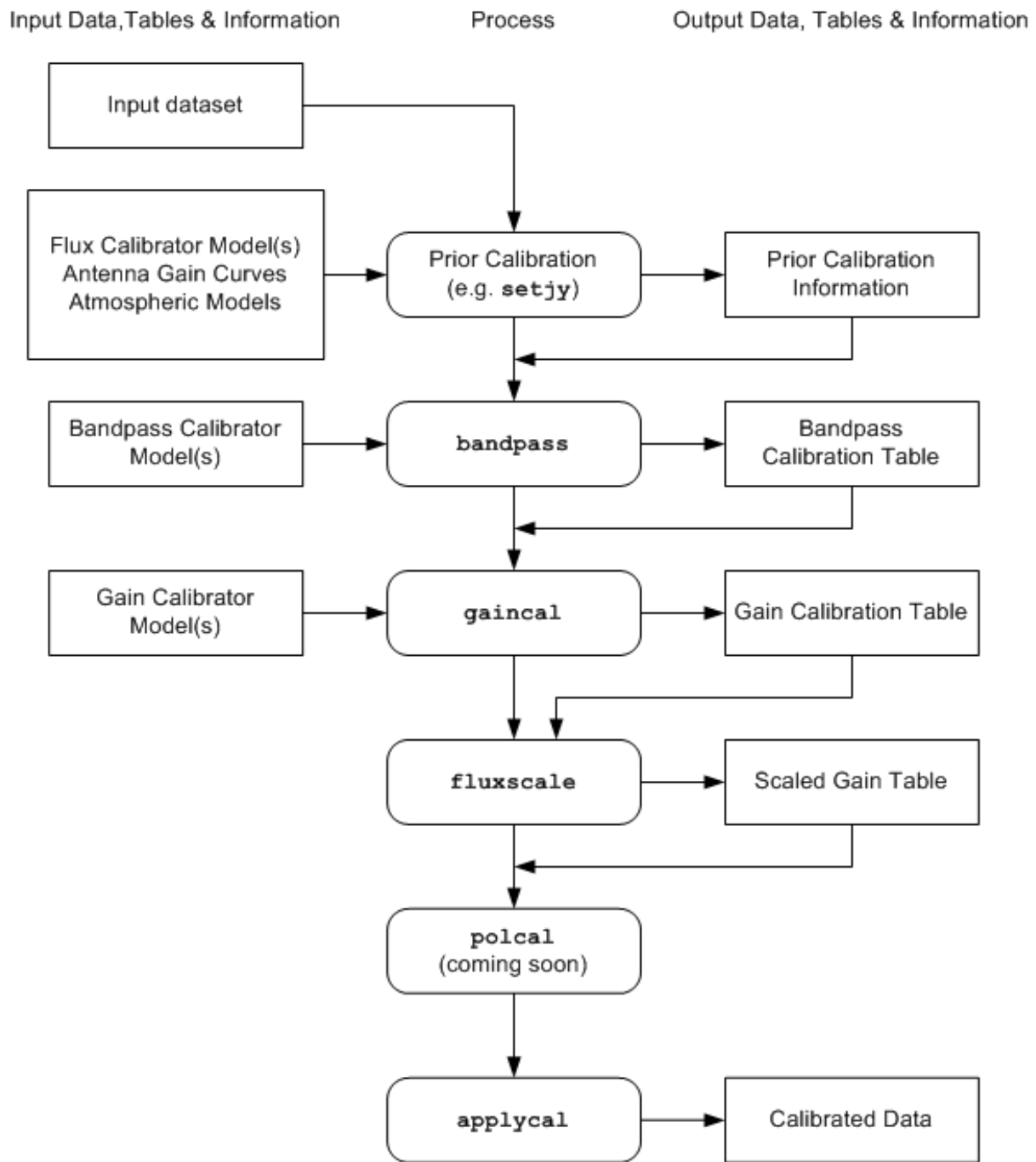


Figure 4.1: Flow chart of synthesis calibration operations. Not shown are use of table manipulation and plotting tasks `accum`, `plotcal`, and `smoothcal` (see Figure 4.2).

- **Calibrator Model Visibility Specification** — set model visibilities for calibrators, either unit point source visibilities for calibrators with unknown flux density or structure (generally, sources used for calibrators are approximately point-like), or visibilities derived from *a priori* images and/or known or standard flux density values.
- **Prior Calibration** — set up previously known calibration quantities that need to be pre-applied, such antenna gain-elevation curves, atmospheric models, delays, and antenna position offsets. Use the `setjy` task (§ 4.3.5) for calibrator flux densities and models, and use `genical` (§ 4.3.6) for antenna position offsets, gaincurves, antenna efficiencies, and opacities;
- **Bandpass Calibration** — solve for the relative gain of the system over the frequency channels in the dataset (if needed), having pre-applied the prior calibration. Use the `bandpass` task (§ 4.4.2);
- **Gain Calibration** — solve for the gain variations of the system as a function of time, having pre-applied the bandpass (if needed) and prior calibration. Use the `gaincal` task (§ 4.4.3);
- **Polarization Calibration** — solve for polarization leakage terms and linear polarization position angle (§ 4.4.5);
- **Establish Flux Density Scale** — if only some of the calibrators have known flux densities, then rescale gain solutions and derive flux densities of secondary calibrators. Use the `fluxscale` task (§ 4.4.4);
- **Manipulate, Accumulate, and Iterate** — if necessary, accumulate different calibration solutions (tables), smooth, and interpolate/extrapolate onto different sources, bands, and times. Use the `accum` (§ 4.5.5) and `smoothcal` (§ 4.5.4) tasks;
- **Examine Calibration** — at any point, you can (and should) use `plotcal` (§ 4.5.1) and/or `listcal` (§ 4.5.2) to look at the calibration tables that you have created;
- **Apply Calibration to the Data** — this can be forced explicitly by using the `applycal` task (§ 4.6.1), and can be undone using `clearcal` (§ 4.6.3);
- **Post-Calibration Activities** — this includes the determination and subtraction of continuum signal from line data, the splitting of data-sets into subsets (usually single-source), and other operations (such as model-fitting). Use the `uvcontsub` (§ 4.7.6), `split` (§ 4.7.1), and `uvmodelfit` (§ 4.7.8) tasks.

The flow chart and the above list are in a suggested order. However, the actual order in which you will carry out these operations is somewhat fluid, and will be determined by the specific data-reduction use cases you are following. For example, you may need to do an initial **Gain Calibration** on your bandpass calibrator before moving to the **Bandpass Calibration** stage. Or perhaps the polarization leakage calibration will be known from prior service observations, and can be applied as a constituent of Prior Calibration.

4.2.1 The Philosophy of Calibration in CASA

Calibration is not an arbitrary process, and there is a methodology that has been developed to carry out synthesis calibration and an algebra to describe the various corruptions that data might be subject to: the Hamaker-Bregman-Sault Measurement Equation (ME), described in Appendix E. The user need not worry about the details of this mathematics as the CASA software does that for you. Anyway, it's just matrix algebra, and your familiar scalar methods of calibration (such as in AIPS) are encompassed in this more general approach.

There are a number of “physical” components to calibration in CASA:

- **data** — in the form of the Measurement Set (§ 2.1). The MS includes a number of columns that can hold calibrated data, model information, and weights;
- **calibration tables** — these are in the form of standard CASA tables, and hold the calibration solutions (or parameterizations thereof);
- **task parameters** — sometimes the calibration information is in the form of CASA task parameters that tell the calibration tasks to turn on or off various features, contain important values (such as flux densities), or list what should be done to the data.

At its most basic level, Calibration in CASA is the process of taking “uncalibrated” **data**, setting up the operation of calibration tasks using **parameters**, solving for new calibration **tables**, and then applying the calibration tables to form “calibrated” **data**. Iteration can occur as necessary, with the insertion of other non-calibration steps (e.g. imaging to generate improved source models for “self-calibration”).

4.2.2 Keeping Track of Calibration Tables

The calibration tables are the currency that is exchanged between the calibration tasks. The “solver” tasks (`gaincal`, `bandpass`, `blcal`, `polcal`) take in the MS (which may have a calibration model attached) and previous calibration tables, and will output an “incremental” calibration table (it is incremental to the previous calibration, if any). This table can then be smoothed using `smoothcal` if desired.

You can optionally accumulate the incremental calibration onto previous calibration tables with `accum`, which will then output a cumulative calibration table. This task will also interpolate onto a different time scale. See § 4.5.5 for more on accumulation and interpolation.

Figure 4.2 graphs the flow of these tables through the sequence

```
solve => smooth => accumulate
```

Note that this sequence applied to separate *types* of tables (e.g. 'B', 'G') although tables of other types can be previous calibration input to the solver.

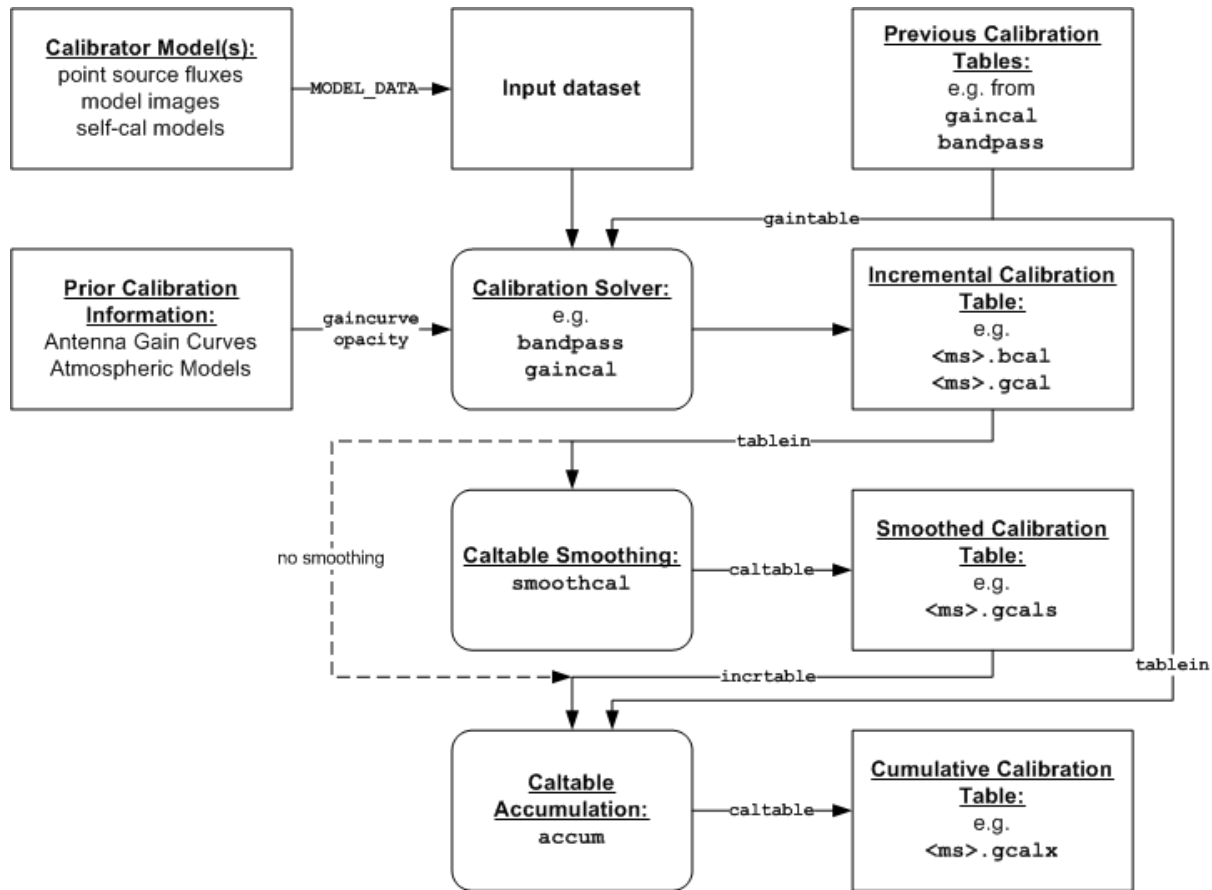


Figure 4.2: Chart of the table flow during calibration. The parameter names for input or output of the tasks are shown on the connectors. Note that from the output solver through the accumulator only a single calibration type (e.g. 'B', 'G') can be smoothed, interpolated or accumulated at a time. `accum` is optional (and not recommended as of v4.0). The final set of cumulative calibration tables of all types (accumulated or as a list of caltables) are then input to `applycal` as shown in Figure 4.1.

The final set of cumulative calibration tables is what is applied to the data using `applycal`. You will have to keep track of which tables are the intermediate incremental tables, and which are cumulative, and which were previous to certain steps so that they can also be previous to later steps until accumulation. This can be a confusing business, and it will help if you adopt a consistent table naming scheme (see Figure 4.2) for an example naming scheme).

4.2.3 The Calibration of traditional VLA data in CASA

CASA supports the calibration of traditional VLA data that is imported from the Archive through the `importvla` task. See § 2.2.3 for more information.

ALERT: Data taken both before and after the Modcomp turn-off in late June 2007 will be handled automatically by `importvla`. You do not need to set special parameters to do so, and it will obey the scaling specified by `applytsys`.

You can also import VLA data in UVFITS format with the `importuvfits` task (§ 2.2.5.1). However, in this case, you must be careful during calibration in that some prior or previous calibrations (see below) may or may not have been done in AIPS and applied (or not) before export.

For example, the default settings of AIPS `FILLM` will apply VLA `gaincurve` and approximate (weather-based) atmospheric optical depth corrections when it generates the extension table `CL 1`. If the data is exported immediately using `FITTP`, then this table is included in the UVFITS file. However, CASA is not able to read or use the AIPS `SN` or `CL` tables, so that prior calibration information is lost and must be applied during calibration here (i.e. using `gaincurve=True` and setting the `opacity` parameter).

On the other hand, if you apply calibration in AIPS by using the `SPLIT` or `SPLAT` tasks to apply the `CL` tables before exporting with `FITTP`, then this calibration will be in the data itself. In this case, you do not want to re-apply these calibrations when processing in CASA.

4.2.4 Loading Jansky VLA data in CASA

Jansky VLA data can be loaded into CASA either via `importevla` or by using the task `importasdm`. Both tasks will convert ASDM raw data files into Measurement Sets. `importasdm` will convert the data itself and the majority of the metadata. `importevla` will run `importasdm` followed by Jansky VLA-specific corrections, like the application of the on-line flags (e.g. times when the subreflector was not in place or the an antenna was not on source), an option to clip values that are exactly zero (as of 2010, such values still may appear in the VLA raw data), and flagging for shadowing.

4.3 Preparing for Calibration

There are a number of “a priori” calibration quantities that may need to be initialized or estimated before further calibration solving is carried out. These include

- **weight initialization** — if desired, initialization of spectral weights (by default, unchannelized weight accounting is used, and no special action is required)
- **system temperature correction** — turn correlation coefficient into correlated flux density (necessary for some telescopes),
- **gain curves** — antenna gain-elevation dependence,
- **atmospheric optical depth** — attenuation of the signal by the atmosphere, including correcting for its elevation dependence.
- **flux density models** — establish the flux density scale using “standard” calibrator sources, with models for resolved calibrators,

- **delays** — antenna-based delay offsets,
- **antenna position errors** — offsets in the positions of antennas assumed during correlation.
- **ionosphere** — dispersive delay and Faraday effects arising from signal transmission through the magnetized plasma of the ionosphere.

These are pre-determined effects and should be applied (if known) as priors when solving for other calibration terms, and included in the final application of all calibration. If unknown, then they will be solved for or subsumed in other calibration such as bandpass or gains.

We now deal with these in turn.

4.3.1 Weight initialization and WEIGHT_SPECTRUM

See Appendix F for a more complete description of weight accounting in CASA.

CASA 4.3 introduced initial experimental support for spectral weights. At this time, this is mainly relevant to ALMA processing for which *spectral* T_{sys} corrections, which faithfully reflect spectral sensitivity, are available. In most other cases, sensitivity is, to a very good approximation, channel-independent after bandpass calibration (and often also before), except perhaps at the very edges of spectral windows (and for which analytic expressions of the sensitivity loss are generally unavailable). Averaging of data with channel-dependent flagging which varies on sufficiently short timescales will also generate channel-dependent net weights (see `split2` or `mstransform` for more details).

By default, CASA’s weight accounting scheme maintains unchannelized weight information that is appropriately updated when calibration is applied. In the case of spectral calibrations (T_{sys} and bandpass), an appropriate spectral average is used for the weight update. This spectral average is formally correct for weight update by bandpass. For T_{sys} , traditional treatments used a single measurement per spectral window; ALMA has implemented spectral T_{sys} to better track sensitivity as a function of channel, and so should benefit from spectral weight accounting as described here, especially where atmospheric emission lines occur. If spectral weight accounting is desired, users must re-initialize the spectral weights using the `initweights` task:

```
initweight(vis='mydata.ms', wtmode='nyq', dowtsp=True)
```

In this task, the `wtmode` parameter controls the weight initialization convention. Usually, when initializing the weight information for a raw dataset, one should choose `wtmode='nyq'` so that the channel bandwidth and integration time information are used to initialize the weight information. The `dowtsp` parameter controls whether (T) or not (F) the spectral weights (WEIGHT.SPECTRUM column) are initialized. The default is `dowtsp=False`, wherein only the non-spectral weights (WEIGHT column) will be initialized. If the spectral weights have been initialized, then downstream processing that supports spectral weights will use/update them. In CASA 4.3 and later, this includes `applycal`, `clean`, and `split2/mstransform`; use of spectral weights in calibration solving (e.g., `gaincal` and other solve tasks) is scheduled for the CASA 4.5 release.

Note that `importasdm` and `importevla` currently initialize the *non-spectral* weights using channel bandwidth and integration time information (equivalent to `initweights(vis='mydata.ms',wtmode='nyq',downtsp=`). In general, it only makes sense to run `initweights` on a raw dataset which has not yet been calibrated, and it should only be necessary if the filled weights are inappropriate, or if spectral weight accounting is desired in subsequent processing. It is usually *not* necessary to re-initialize the weight information when redoing calibration from scratch (the raw weight information is preserved in the SIGMA/SIGMA.SPECTRUM columns). (Re-)initializing the weight information for data that has already been calibrated (with `calwt=T`, presumably) is formally incorrect and is not recommended.

When combining datasets from different epochs, it is generally preferable to have used the same version of CASA (most recent is best), and with the same weight information conventions and `calwt` settings. Doing so will minimize the likelihood of arbitrary weight imbalances that might lead to net loss of sensitivity, and maximize the likelihood that *real* differences in per-epoch sensitivity (e.g., due to different weather conditions and instrumental setups) will be properly accounted for. Modern instruments support more variety in bandwidth and integration time settings, and so use of these parameters in weight initialization is preferred (c.f. use of unit weight initialization, which has often been the traditional practice).

Alert: Full and proper weight accounting for the EVLA formally depends on the veracity of the switched power calibration scheme (§ 4.3.7). As of mid-2015, use of the EVLA switched power is not yet recommended for general use, and otherwise uniform weights are carried through the calibration process. As such, spectral weight accounting is not yet meaningful. Facilities for post-calibration estimation of spectral weights are planned for a future release.

4.3.2 System Temperature and Switched-Power Corrections

Some telescopes, including the old VLA, ALMA, and the VLBA, record the visibilities in the form of raw *correlation coefficient* with weights proportional to the number of bits correlated. The correlation coefficient is the fraction of the total signal that is correlated, and thus multiplication by the system temperature (T_{sys}) and the antenna gain (in Jy/K) will produce visibilities with units of correlated flux density. ALMA records $T_{sys}(K)$ information in the MS which can be extracted as a caltable using `genical` with `calmode='tsys'`, and applied to data to yield units of K . Calibration to flux density in Jy is achieved via reference to sources of known power.

Alert: Note that the old VLA system did this initial calibration on-line. The modern VLA does not record normalized visibilities. Instead, the correlations are delivered in raw engineering units that are proportional to power. The actual total power received is continuously monitored during the observation, with a calibration signal of known temperature (K) switched in at a rate of 10 Hz. This is the so-called “switched-power” calibration system on the VLA. This enables a continuous record of the $T_{sys}(K)$, as well as net electronic gain variation of each antenna’s receiving system. The correlator requantizer gain is also monitored. These data are recorded in MS subtables and appropriate calibration factors can be derived from them by `genical` with `caltype='swpow'`, and stored in a caltable for application. This calibration is *not* a “ T_{sys} ” calibration of the traditional sort; the switched-power gain is used to correct the visibility amplitude, and the T_{sys} is used to set the weights. This system is still being commissioned (as of early 2014). Observations using 8-bit sampling are usually reasonably calibrated; 3-bit-sampled switched-power data are

subject to compression effects that are not yet completely understood, and the switched power calibration is not recommended (instead, correction only by the requantizer gain is recommended, using `caltype='rq'`).

See § 4.3.6 for more information on use of `gencal`.

4.3.3 Antenna Gain-Elevation Curve Calibration

Large antennas (such as the 25-meter antennas used in the VLA and VLBA) have a forward gain and efficiency that changes with elevation. Gain curve calibration involves compensating for the effects of elevation on the amplitude of the received signals at each antenna. Antennas are not absolutely rigid, and so their effective collecting area and net surface accuracy vary with elevation as gravity deforms the surface. This calibration is especially important at higher frequencies where the deformations represent a greater fraction of the observing wavelength. By design, this effect is usually minimized (i.e., gain maximized) for elevations between 45 and 60 degrees, with the gain decreasing at higher and lower elevations. Gain curves are most often described as 2nd- or 3rd-order polynomials in zenith angle.

Gain curve calibration has been implemented in CASA for the modern VLA and old VLA (only), with gain curve polynomial coefficients available directly from the CASA data repository. To make gain curve and antenna efficiency corrections for VLA data, use `gencal` with `caltable='gceff'`. See § 4.3.6 for more information on use of `gencal`.

ALERT: If you are not using VLA data, do not use `gaincurve` corrections. A general mechanism for incorporating `gaincurve` information for other arrays will be made available in future releases. The gain-curve information available for the VLA is time-dependent (on timescales of months to years, at least for the higher frequencies), and CASA will automatically select the date-appropriate gain curve information. Note, however, that the time-dependence was poorly sampled prior to 2001, and so gain curve corrections prior to this time should be considered with caution.

4.3.4 Atmospheric Optical Depth Correction

The troposphere is not completely transparent. At high radio frequencies (>15 GHz), water vapor and molecular oxygen begin to have a substantial effect on radio observations. According to the physics of radiative transmission, the effect is threefold. First, radio waves from astronomical sources are absorbed (and therefore attenuated) before reaching the antenna. Second, since a good absorber is also a good emitter, significant noise-like power will be added to the overall system noise. Finally, the optical path length through the troposphere introduces a time-dependent phase error. In all cases, the effects become worse at lower elevations due to the increased air mass through which the antenna is looking. In CASA, the opacity correction described here compensates only for the first of these effects, tropospheric attenuation, using a plane-parallel approximation for the troposphere to estimate the elevation dependence.

To make opacity corrections in CASA, an estimate of the zenith opacity is required (see observatory-specific chapters for how to measure zenith opacity). This is then supplied to the `caltype='opac'`

parameter in `gencal` which creates a calibration table with all the information. E.g. for data with two spectral windows, the inputs are like:

```
gencal(vis='dataset.ms',
       caltable='opacity.cal',
       caltype='opac',
       spw='0,1',
       parameter=[0.0399,0.037])
```

If you do not have an externally supplied value for `opacity`, for example from a VLA tip procedure, then you should either use an average value for the telescope, or leave it at zero and let your gain calibration compensate as best it can (e.g. that your calibrator is at the same elevation as your target at approximately the same time. As noted above, there are no facilities yet to estimate this from the data (e.g. by plotting T_{sys} vs. elevation).

Below, we give instructions for determining `opacity` for Jansky VLA data from weather statistics and VLA observations where tip-curve data is available. It is beyond the scope of this cookbook to provide information for other telescopes.

4.3.4.1 Determining opacity corrections for *modern VLA* data

For the VLA site, weather statistics and/or seasonal models that average over many years of weather statistics prove to be reasonable good ways to estimate the opacity at the time of the observations. The task `plotweather` calculates the opacity as a mix of both actual weather data and seasonal model. It has the following inputs:

```
# plotweather :: Plot elements of the weather table; estimate opacity.
vis           =      ''           # MS name
seasonal_weight =      0.5       # weight of the seasonal model
doPlot       =      True        # set this to True to create a plot
```

The task plots the weather statistics if `doPlot=T`, like shown in Figure 4.3. The bottom panel displays the calculated opacities for the run as well as a seasonal model. The parameter `seasonal_weight` can be adjusted to calculate the opacities as a function of the weather data alone `seasonal_weight=0`, only the seasonal model `seasonal_weight=1`, or a mix of the two (values between 0 and 1). Calculated opacities are shown in the logger output, one for each spectral window. `plotweather` can also assign a python variable to a list of calculated opacities (one entry for each spw) when run as:

```
myTau = plotweather(vis='myvladata.ms')
```

In this example, `myTau` will be returned with a list of per-spw opacities, e.g. `myTau=[0.02,0.03]` and can later be used as input for `gencal` in `caltype='opac'` in the `parameter` setting, e.g.,

```
# opac for spws 0,1 in myTau
gencal(vis='myvladata.ms',caltype='opac',spw='0,1',parameter=myTau)
```

Note that it is important to explicitly specify the spws that are covered by the opacity values stored in `myTau`. For most modern VLA data there will be more than two spws, probably.

See § 4.3.6 for more information on use of `gencal`.

4.3.4.2 Determining opacity corrections for VLA data

For VLA data, zenith opacity can be measured at the frequency and during the time observations are made using a VLA tipping scan in the observe file. Historical tipping data are available at:

<http://www.vla.nrao.edu/astro/calib/tipper>

Choose a year, and click **Go** to get a list of all tipping scans that have been made for that year.

If a tipping scan was made for your observation, then select the appropriate file. Go to the bottom of the page and click on the button that says **Press here to continue..** The results of the tipping scan will be displayed. Go to the section called 'Overall Fit Summary' to find the fit quality and the fitted zenith opacity in percent. If the zenith opacity is reported as 6%, then the actual zenith optical depth value is `opacity=0.060` for `gaincal` and other calibration tasks.

If there were no tipping scans made for your observation, then look for others made in the same band around the same time and weather conditions. If nothing is available here, then at K and Q bands you might consider using an average value (e.g. 6% in reasonable weather). See the VLA memo

<http://www.vla.nrao.edu/memos/test/232/232.pdf>

for more on the atmospheric optical depth correction at the VLA, including plots of the seasonal variations.

4.3.5 Setting the Flux Density Scale using (`setjy`)

When solving for visibility-plane calibration, CASA calibration applications compare the observed `DATA` column with the Fourier transform of calibrator model when it is provided (if no model is specified, a point source at the phase center is assumed).

The `setjy` task is used to set the proper flux density and attaches a model image (if specified) of the calibrator to the MS. For sources that are recognized flux calibrators (listed in Tables 4.1 and 4.2, see also § C.1), `setjy` can calculate the flux densities as a function of frequency (and time, for Solar System objects). Otherwise, the flux densities should be manually specified (`standard='manual'`).

For the VLA, the default source models are customarily point sources defined by the 'Baars', 'Perley 90', 'Perley-Taylor 99', 'Perley-Butler 2010', time-variable 'Perley-Butler 2013', or 'Scaife-Heald 2012' flux density scales (§ C.1.1; 'Perley-Butler 2013' is the current standard by default), or point sources of unit flux density if the flux density is unknown. In fact, the model can be any image in Jy/pixel units (models typically generated by the `clean` task).

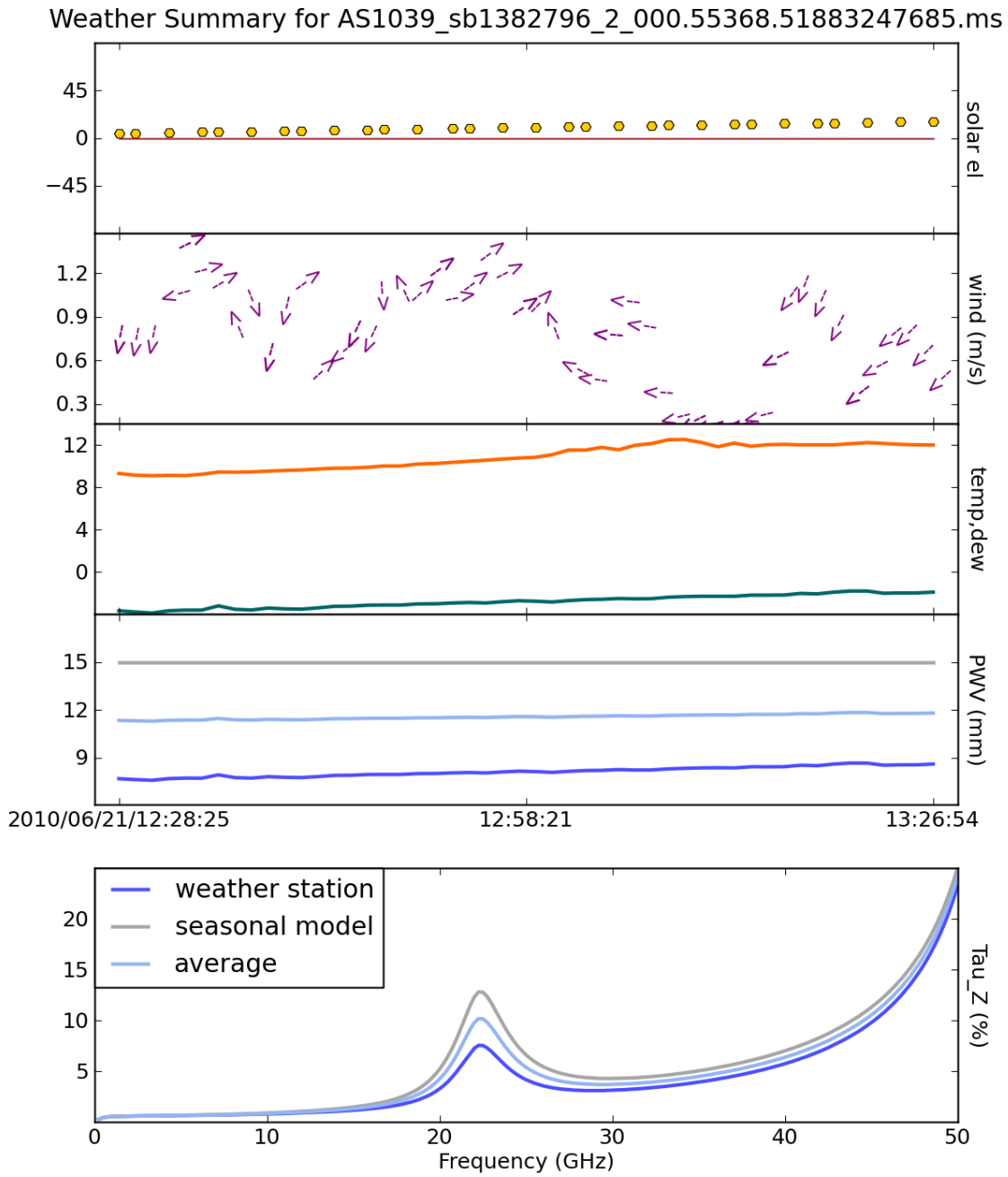


Figure 4.3: The weather information for a MS as plotted by the task `plotweather`.

Optionally, the MODEL column can be filled with the Fourier transform of (option `usescratch=T` is `setjy`, `ft`, and `clean`). But for most Measurement Sets, the performance and data storage requirements are less demanding without the MODEL_DATA column.

The inputs for `setjy` are:

```
# setjy :: Fills the model column with the visibilities of a calibrator
vis          =      ''          # Name of input visibility file
field        =      ''          # Field name(s)
spw          =      'all'       # Spectral window identifier (list)
selectdata   =      True        # Other data selection parameters
  timerange  =      ''          # Time range to operate on (for usescratch=T)
  scan       =      ''          # Scan number range (for usescratch=T)
  intent     =      ''          # Observation intent
  observation =      ''          # Observation ID range (for usescratch=T)

scalebychan  =      True        # scale the flux density on a per channel basis or else on
                                # a per spw basis
standard     =      'Perley-Butler 2013' # Flux density standard
  model      =      ''          # File location for field model
  listmodels =      False       # List the available modimages for VLA calibrators or Tb
                                # models for Solar System objects

usescratch   =      False       # Will create if necessary and use the MODEL_DATA
```

Table 4.1: Recognized Flux Density Calibrators. Note that the VLA uses J2000 calibrator names. CASA accepts all strings that contain the names below. E.g. 'PKS 1934-638' will be recognized

3C Name	B1950 Name	J2000 Name	Alt. J2000 Name	Standards
3C48	0134+329	0137+331	J0137+3309	1,3,4,5,6,7
3C123	0433+295	0437+296	J0437+2940	2
3C138	0518+165	0521+166	J0521+1638	1,3,4,5,6
3C147	0538+498	0542+498	J0542+4951	1,3,4,5,6,7
3C196	0809+483	0813+482	J0813+4813	1,2,7
3C286	1328+307	1331+305	J1331+3030	1,2,3,4,5,6,7
3C295	1409+524	1411+522	J1411+5212	1,2,3,4,5,6,7
–	1934-638	–	J1939-6342	1,3,4,5,6
3C380	1828+487	1829+487	J1829+4845	7

Standards are: (1) Perley-Butler 2010, (2) Perley-Butler 2013, (3) Perley-Taylor 99, (4) Perley-Taylor 95, (5) Perley 90, (6) Baars (Baars, J. W. M., et al. 1977, A&A, 61, 99); (7) Scaife-Heald 2012, see § C.1.1 for details.

By default the `setjy` task will cycle through all fields spectral windows and channels, (one solution per spw with `scalebychan = False`), setting the flux density either to 1 Jy (unpolarized), or if the source is recognized as one of the calibrators in the above table, to the flux density (assumed

Table 4.2: 'Butler-JPL-Horizons 2012' recognized Solar System Objects for Flux Calibration

Planets Venus ¹ , Mars ² , Jupiter ³ , Uranus ⁴ , Neptune ⁵
Moons <i>Jupiter</i> : Io, Europa, Ganymede, Callisto <i>Saturn</i> : Titan ⁷
Asteroids Ceres, Pallas ⁸ , Vesta ⁸ , Juno ⁸

¹ Venus: model for ~ 300 MHz to 350 GHz, no atmospheric lines (CO, H₂O, HDO, etc.)

² Mars: tabulated as a function of time and frequency (30 - 1000GHz) based on Rudy et al. (1988), no atmospheric lines (CO, H₂O, H₂O₂, HDO, etc.)

³ Jupiter: model for 30-1020GHz, does not include synchrotron emission

⁴ Uranus: model for 60-1800GHz, contains no rings or synchrotron.

⁵ Neptune: model for 2-2000GHz, the broad CO absorption line is included, but contains no rings or synchrotron.

⁷ Titan: model for 53.3-1024.1GHz, include many spectral lines

⁸ not recommended (The temperature is not yet adjusted for varying distance from the Sun. The model data can be scaled after running `setjy`, but it is an involved process.)

Details are described in ALMA Memo 594 available on https://science.nrao.edu/facilities/alma/aboutALMA/Technology/ALMA_Memo_Series/alma594/abs594.

unpolarized) appropriate to the observing frequency. For example, to run `setjy` on a Measurement Set called `data.ms`:

```
setjy(vis='data.ms')           # This will set all fields and spectral windows
```

Models of available calibrator sources can be listed by setting `listmodels=True`. `setjy` will then come up with all images that are in the paths where calibrator models for known telescopes are stored. It will also show all images in the working directory - any image there could potentially be a calibrator model. If the calibrator model is found by `listmodels` it can be used in the `modimage` parameter without a path.

The `fluxdensity` parameter can be used to specify the flux of the calibrator in all Stokes parameters. It is thus a list of values [I,Q,U,V], e.g. ['12Jy','13mJy','0Jy','0Jy']. In addition, a spectral index can be specified via `spix`, a reference frequency `reffreq` (using the definition: $S = \text{fluxdensity} \times \frac{\text{freq}^{\text{spix}}}{\text{reffreq}}$), as well as a polarization index (`polindex`), angle (`polangle`) and a rotation measure (`rotmeas`).

Most calibrator sources are based on radio emission from quasars and jets. The spectral indices of these sources are such that at (sub)mm wavelengths the majority of these sources become too weak and variable to be reliable flux estimators. Alternatives are thermal objects such as planets, moons, and asteroids. Those sources, however, are all Solar System objects, which implies that they move and may be (strongly) resolved. The `standard='Butler-JPL-Horizons 2010'` and the recommended `standard='Butler-JPL-Horizons 2012'` (for more information on the implemented models, see ALMA Memo 594 soon available on <https://science.nrao.edu/facilities/alma/>

aboutALMA/Technology/ALMA_Memo_Series/alma594/abs594.) option of `setjy` includes flux density calibration using Solar System objects. For 'Butler-JPL-Horizons 2012' CASA currently supports the objects listed in Table 4.2 to be applied to ALMA data. These names are recognized when they are used in the 'field' parameter in `setjy`. In that case, `setjy` will obtain the geocentric distance and angular diameter at the time of the observation from a (JPL-Horizons) ephemeris and calculate model visibilities. Currently the objects are modeled as uniform temperature disks, but effects like primary beam attenuation and limb darkening will be accounted for soon. Note that this model may oversimplify the real structure, in particular asteroids.

An example, using data from the M99 tutorial in http://casaguides.nrao.edu/index.php?title=CARMA_spectral_line_mosaic_M99:

```
setjy(vis='c0104I', field='MARS', spw='0 2', standard='Butler-JPL-Horizons 2012')
```

Tip: Running `casalog.filter('INFO1')` before running `setjy` with a Solar System object may send the logger a reference to the temperature measurement. Use `casalog.filter('INFO')` to restore the normal logging level.

The source model will be attached to the MS and applied to all calibration steps when `usescratch=False`. `usescratch=True` fills the MODEL_DATA column with the Fourier transform of the model. As of CASA 3.4, we found that under some circumstances, creation of the MODEL column may prevent memory issues and if tasks fail, we recommend to set `usescratch=True`. Note that currently `setjy` will not transform a full-Stokes model image such that all polarizations are applied correctly. You need to use `ft` for this.

To limit this operation to certain fields and spectral windows, use the `field` and/or `spw` parameters, which take the usual data selection strings (§ 2.3). For example, to set the flux density of the first field (all spectral windows)

```
setjy(vis='data.ms',field='0')
```

or to set the flux density of the second field in spectral window 17

```
setjy(vis='data.ms',field='1',spw='17')
```

The full-polarization flux density (I,Q,U,V) may also be explicitly provided:

```
setjy(vis='data.ms',
      field='1',spw='16',           # Run setjy on field id 1, spw id 17
      fluxdensity=[3.5,0.2,0.13,0.0]) # and set I,Q,U,V explicitly
```

ALERT: The apparent brightness of objects in the Solar System will vary with time because of the Earth's varying distance to them, if nothing else. If the field *index* of a flux calibrator spans several days, you should run `setjy` more than once, limiting each run to a suitable timerange by using the `timerange`, `scan`, and/or observation selection parameters. Note that it is the field index that matters, not the name. Typically `concat` assigns moving objects a new field index for each observation, so usually it is not necessary to select a time range in `setjy`. However, it is worth checking with `listobs`, especially for planets.

4.3.5.1 Using Calibration Models for Resolved Sources

For observations of solar system objects using the 'Butler-JPL-Horizons 2010' and 'Butler-JPL-Horizons 2012' models (§ 4.3.5) `setjy` will know and apply the flux distribution across the extended structure of the calibrators.

For other sources, namely VLA calibrator sources, a flux density calibrator can be resolved at the observing frequency and the point source model generated by `setjy` will not be appropriate. If available, a model image of the resolved source at the observing frequency may be used to generate the appropriate visibilities using the `modimage` parameter (or in older versions explicitly with the `ft` task). To use this, provide `modimage` with the path to the model image. Remember, if you just give the file name, it will assume that it is in the current working directory. Note also that `setjy` using a model image will only operate on that single source, thus you would run it multiple times (with different `field` settings) for different sources.

Otherwise, you may need to use the `uvrange` selection (§ 4.4.1.2) in the calibration solving tasks to exclude the baselines where the resolution effect is significant. There is not hard and fast rule for this, though you should consider this if your calibrator is shows a drop of more than 10% on the longest baselines (use `plotxy`, § 3.3.2, to look at this). You may need to do `antenna` selection also, if it is heavily resolved and there are few good baselines to the outer antennas. Note that `uvrange` may also be needed to exclude the short baselines on some calibrators that have extended flux not accounted for in the model. **Note:** the calibrator guides for the specific telescopes usually indicate appropriate min and max for `uvrange`. For example, see the *VLA Calibration Manual* at:

<http://www.vla.nrao.edu/astro/calib/manual/>

for details on the use of standard calibrators for the E/VLA.

Model images for some flux density calibrators are provided with CASA:

- Red Hat Linux RPMs 32bit (RHE4, Fedora 6): located in
/usr/lib/casapy/data/nrao/VLA/CalModels
- Red Hat Linux RPMs 64bit (RHE4, Fedora 6): located in
/usr/lib64/casapy/data/nrao/VLA/CalModels
- MAC OSX .dmg: located in
/Applications/CASA.app/Contents/Resources/casa-data/nrao/VLA/CalModels
- NRAO-AOC casapy-test:
/home/casa/data/nrao/VLA/CalModels

e.g., these are found in the `data/nrao/VLA/CalModels` sub-directory of the CASA installation. For example, just point to the repository copy, e.g.

```
modimage = '/usr/lib/casapy/data/nrao/VLA/CalModels/3C48_C.im'
```

or if you like, you can copy the ones you wish to use to your working directory.

The models available are:

```
3C138_L.im 3C147_L.im 3C286_L.im 3C48_L.im
3C138_C.im 3C147_C.im 3C286_C.im 3C48_C.im
3C138_X.im 3C147_X.im 3C286_X.im 3C48_X.im
3C138_U.im 3C147_U.im 3C286_U.im 3C48_U.im
3C138_K.im 3C147_K.im 3C286_K.im 3C48_K.im
3C138_Q.im 3C147_Q.im 3C286_Q.im 3C48_Q.im
```

(more calibrator models for the VLA are available at <https://science.nrao.edu/facilities/vla/data-processing/models> These are all un-reconvolved images of AIPS CC lists. It is important that the model image *not* be one convolved with a finite beam; it must have units of Jy/pixel (not Jy/beam).

Note that `setjy` will rescale the flux in the models for known sources (e.g. those in Table 4.1) to match those it would have calculated. It will thus extrapolated the flux out of the frequency band of the model image to whatever spectral windows in the MS are specified (but will use the structure of the source in the model image).

ALERT: The reference position in the `modimage` is currently used by `setjy` when it does the Fourier transform, thus differences from the positions for the calibrator in the MS will show up as phase gradients in the uv-plane. If your model image position is significantly different but you don't want this to affect your calibration, then you can doctor either the image header using `imhead` (§ 6.2) or in the MS (using the `ms` tool) as appropriate. In an upcoming release we will put in a toggle to use or ignore the position of the `modimage`. Note that this will not affect the flux scaling (only put in erroneous model phases); in any event small position differences, such as those arising by changing epoch from B1950 to J2000 using `regridimage` (§ 6.13), will be inconsequential to the calibration.

This illustrates the use of `uvrange` for a slightly resolved calibrator:

```
# Import the data
importvla(archivefiles='AS776_A031015.xp2', vis='ngc7538_XBAND.ms',
          freqtol=10000000.0, bandname='X')

# Flag the ACs
flagautocorr('ngc7538_XBAND.ms')

# METHOD 1: Use point source model for 3C48, plus uvrange in solve

# Use point source model for 3C48
setjy(vis='ngc7538_XBAND.ms', field='0');

# Limit 3C48 (fieldid=0) solutions to uvrange = 0-40 klambda
gaincal(vis='ngc7538_XBAND.ms', caltable='cal.G', field='0',
        solint=60.0, refant='10', selectdata=True, uvrange='0~40klambda',
        append=False)
```

```
# Append phase-calibrator's solutions (no uvrage) to the same table
gaincal(vis='ngc7538_XBAND.ms', caltable='cal.G', field='2',
        solint=60.0, refant='10', selectdata=True, uvrage='',
        append=True)

# Fluxscale
fluxscale(vis='ngc7538_XBAND.ms', caltable='cal.G', reference=['0137+331'],
          transfer=['2230+697'], fluxtable='cal.Gflx', append=False)
```

while the following illustrates the use of a model:

```
# METHOD 2: use a resolved model copied from the data repository
#   for 3C48, and no uvrage
# (NB: detailed freq-dep flux scaling TBD)

# Copy the model image 3C48_X.im to the working directory first!

setjy(vis='ngc7538_XBAND.ms', field='0', modimage='3C48_X.im')

# Solutions on both calibrators with no uvrage
gaincal(vis='ngc7538_XBAND.ms', caltable='cal.G2', field='0,2',
        solint=60.0, refant='10',
        append=False)

# Fluxscale
fluxscale(vis='ngc7538_XBAND.ms', caltable='cal.G2', reference=['0137+331'],
          transfer=['2230+697'], fluxtable='cal.G2flx', append=False)

# Both methods give 2230 flux densities ~0.7 Jy, in good agreement with
#   AIPS
```

4.3.6 Correction for delay and antenna position offsets using gencal

The `gencal` task provides a means of specifying antenna-based calibration values manually. The values are put in designated tables and can be applied to the data on-the-fly in solving tasks and using `applycal`.

The `gencal` task has the inputs:

```
# gencal :: Specify Calibration Values of Various Types
vis           =      ''      # Name of input visibility file
caltable      =      ''      # The new/existing calibration table
caltype       =      'tecim'  # The calibration type: 'amp', 'ph', 'sbd', 'mbd', 'antpos', 'an
# tposvla', 'tsys', 'evlagain', 'opac', 'gc', 'gceff', 'eff', 'te
# cim'
infile        =      ''      # Input ancilliary file
spw           =      'all'    # Calibration spw(s) selection
```

```

antenna      =      '1'      # Calibration antenna(s) selection
pol          =      ''       # Calibration polarizations(s) selection
parameter    =      []       # The calibration values

```

Current antenna-based gencal options (caltypes) are:

- 'amp' — amplitude correction
- 'ph' — phase correction
- 'sbd' — single-band delay (phase-frequency slope for each spw)
- 'mbd' — multi-band delay (phase-frequency slope over all spw)
- 'antpos' — ITRF antenna position corrections for the Jansky VLA (automatic parameter lookup is supported)
- 'antposvla' — old VLA-centric antenna position corrections
- 'tsys' — Tsys from the SYSCAL table (ALMA)
- 'evlagain' — VLA switched-power gains (experimental; equal to 'swpow')
- 'swpow' — VLA switched power (equal to 'evlagain')
- 'rq' — VLA requantizer gains *only*
- 'swp/rq' — VLA switched-power gains divided by requantizer gain
- 'opac' — Tropospheric opacity
- 'gc' — Gain curve (zenith-angle-dependent gain) (VLA only) (auto-lookup only)
- 'gceff' — Antenna efficiency (sqrt(K/Jy)) (VLA only) (auto-lookup only)
- 'eff' — Gain curve and efficiency (VLA only) (auto-lookup only)
- 'tecim' — Total electron content to derive dispersive delays

The calibration parameter specifications cannot be time-variable in the present implementation (though some of them will introduce implicit time-dependence upon evaluation in the apply). Calibration values can be assigned to each `spw`, `antenna` and `pol` selection, where applicable. The list of calibration values specified in `parameter` must conform to the range of spectral windows, antennas, and polarizations specified in `spw`, `antenna` and `pol`, with the values specified in order of the specified polarizations (fastest), antennas, and spectral windows (slowest). If any of `spw`, `antenna`, or `pol` are left unspecified (empty strings), the values specified in `parameter` will be assumed applicable to all values of the unspecified data axes. The output caltable will otherwise assume nominal calibration values for unspecified spectral windows, antennas, and polarizations. Note that antenna position corrections formally do not have spectral-window or polarization dependence; such specifications should not be used with 'antpos'.

The same caltable can be specified for multiple runs of `gencal`, in which case the specified parameters will be incorporated cumulatively. E.g., amplitude parameters (`caltype='amp'`) multiply and phase-like parameters (`'ph'`, `'sbd'`, `'mbd'`, `'antpos'`) add. Parameters for `'amp'` and `'ph'` corrections can be incorporated into the same caltable (in separate runs), but each of the other types require their own unique caltable. A mechanism for specifying manual corrections via a text file will be provided in the future.

Two kinds of delay corrections are supported. For `caltype='sbd'`, the specified delays (in nanoseconds) will be applied locally to each spectral window, referring the derived phase corrections to each spectral window's reference frequency (where the phase correction will be zero). The phases in each spectral window will nominally be flattened, but any phase offsets between spectral windows will remain. (These can be corrected using `caltype='phase'`, or via ordinary spectral-window-dependent phase calibration.) For `caltype='mbd'`, the evaluated phase corrections are referred to zero frequency. This causes a correction that is coherent over many spectral windows. If the data are already coherent over many spectral windows and share a common multi-band delay (e.g., VLA data, per baseband), `caltype='mbd'` corrections will maintain this coherence and flatten the frequency-dependent phase. Using `caltype='sbd'` in this instance will introduce phase offsets among spectral windows that reflect the multi-band delay.

For antenna position corrections (`caltype='antpos'`), the antenna position offsets are specified in the ITRF frame. If the `antenna` field is left empty, `gencal` will try to look up the appropriate antenna position offsets at the time of the observation from the VLA baseline webpage <http://www.vla.nrao.edu/astro/archive/baselines/>.

For VLA position corrections in the VLA-centric frame, use `caltype='antposvla'`, and `gencal` will rotate them to ITRF before storing them in the output caltable.

The sign and scale convention for `gencal` corrections (indeed for all CASA caltables) is such that the specified parameters (and as stored in caltables) are the factors that *corrupt* ideal data to yield the observed data. Thus, when applied to correct the data, their effective inverse will automatically be taken. I.e., amplitude factors will be divided into the data on correction. Phase-like parameters adopt the convention that the complex factor for the second antenna in the baseline is conjugated, and then both antenna factors are divided into the data on correction. (These conventions differ from AIPS in that **multiplying** correction factors are stored in AIPS calibration tables; however, the phase convention ends up being the same since AIPS conjugates the complex factor for the *first* antenna in the baseline.)

The following series of examples illustrate the use of `gencal`.

For the dataset `'data.ms'`, the following sequence of `gencal` runs introduces, into a single caltable (`'test.G'`), (1) an antenna-based amplitude scale correction of 3.0 for all polarizations, antennas, and spectral windows, (2) phase corrections for all spectral windows and polarizations of 45 and 120 degrees to antennas EA03 and EA04, respectively, (3) phase corrections for all spectral windows of 63 and -34 in R (only) for antennas EA05 and EA06, respectively, and (4) phase corrections for all spectral windows of 14, -23, -130, and 145 degrees for antenna/polarizations EA09/R, EA09/L, EA10/R, and EA10/L, respectively:

```
gencal(vis='data.ms',caltable='test.G',caltype='amp', \
```

```

    spw='',antenna='',pol='', \
    parameter=[3])

gencal(vis='data.ms',caltable='test.G',caltype='ph', \
    spw='',antenna='EA03,EA04',pol='', \
    parameter=[45,120])

gencal(vis='data.ms',caltable='test.G',caltype='ph', \
    spw='',antenna='EA05,EA06',pol='R', \
    parameter=[63,-34])

gencal(vis='data.ms',caltable='test.G',caltype='ph', \
    spw='',antenna='EA09,EA10',pol='R,L', \
    parameter=[14,-23,-130,145])

```

In the following example, delay corrections in both polarizations will be adjusted for antenna EA09 by 14 nsec in spw 2 and -130 nsec in spw 3, and for antenna EA10 by -23 nsec in spw 2 and 145 nsec in spw 3:

```

gencal(vis='test.ms',caltable='test.sbd',caltype='sbd', \
    spw='2,3',antenna='EA09,EA10',pol='', \
    parameter=[14,-23,-130,145])

```

In the following example, antenna position corrections in meters (in ITRF) for antenna EA09 (dBx=0.01, dBy=0.02, dBz=0.03) and for antenna EA10 (dBx=-0.03, dBy=-0.01, dBz=-0.02) are introduced. Note that three parameters are required for each antenna. The antenna offsets can be obtained for the 'Jansky VLA/ old VLA Baseline Corrections' web page: <http://www.vla.nrao.edu/astro/archive/bas>. The table given on this webpage has a format like:

```

;                2010 BASELINE CORRECTIONS IN METERS
;ANT
;MOVED OBSDATE   Put_In_ MC(IAT)  ANT  PAD   Bx      By      Bz
;
JAN27  FEB12     FEB21  01:57   11   E04  0.0000  0.0000  0.0000
JAN27  FEB12     FEB21  01:57   26   W03 -0.0170  0.0204  0.0041
MAR24  MAR25     MAR26  18:28   17   W07 -0.0061 -0.0069 -0.0055
APR21  MAY02     MAY04  23:25   12   E08 -0.0072  0.0045 -0.0017

```

If your observations fall in between the 'Antenna Moved' and 'Put_In_' dates of a given antenna, you may choose to apply the offsets in that table; the 'Put_In_' time stamp marks the date where the more accurate solution was introduced in the data stream directly and no correction is required anymore. In `gencal` the offsets will be inserted as:

```

gencal(vis='test.ms',caltable='test.antpos',caltype='antpos', \
    antenna='EA09,EA10', \
    parameter=[0.01,0.02,0.03, -0.03,-0.01,-0.02])

```


In the following example, antenna position corrections (in the traditional VLA-centric frame) will be introduced in meters for antenna EA09 (dBx=0.01, dBy=0.02, dBz=0.03) and for antenna EA10 (dBx=-0.03, dBy=-0.01, dBz=-0.02) These offsets will be rotated to the ITRF frame before storing them in the caltable.

```
gencal(vis='test.ms', caltable='test.antposvla', caltype='antposvla', \
       antenna='EA09,EA10', \
       parameter=[0.01,0.02,0.03, -0.03,-0.01,-0.02])
```

`gencal` is also the task to generate gaincurve, antenna efficiency, and opacity tables. The first two items can be determined together with `caltype='gceff'` and the latter with `caltype='opac'`. These tables are treated just like any other calibration table and will be carried through the calibration steps. This method replaces the older method where 'gaincurve' and 'opacity' keywords were present in calibration tasks such as `gaincal`, `bandpass`, or `applycal`.

4.3.7 Applying Jansky VLA switched power or ALMA Tsys using gencal

Noise diodes in the Jansky VLA antennas can be used to pre-calibrate the data. The diodes follow an ON-OFF cycle and the power for both states is measured and recorded. This is called the 'VLA switched power' calibration. To apply the switched power data, one needs to create a calibration table with `gencal` using `caltype='evlagain'`, like

```
gencal(vis='test.ms', caltable='VLAswitchedpower.cal', caltype='evlagain')
```

For ALMA the calibration of system temperature is done via hot loads and the data recorded similar to the VLA in the Measurement Set (ALMA will provide Measurement Sets where these data are available. To derive the calibration table from it, use `caltype='tsys'`:

```
gencal(vis='test.ms', caltable='ALMAtsys.cal', caltype='tsys')
```

This calibration tables created for ALMA or VLA are then carried along all further calibration steps in the `gaintable` parameter.

4.3.8 Generate a gain table based on Water Vapor Radiometer data wvrgcal

```
# wvrgcal :: Generate a gain table based on Water Vapour Radiometer data
vis           =      ''      # Name of input visibility file
caltable      =      ''      # Name of output gain calibration table
toffset      =      0        # Time offset (sec) between interferometric and WVR data
segsource    =      True     # Do a new coefficient calculation for each source
    tie       =      []      # Prioritise tying the phase of these sources as well as
    # possible (requires segsource=True)
    sourceflag =      []      # Flag the WVR data for these source(s) as bad and do not
```

```

# produce corrections for it (requires segsource=True)
disperse = False # Apply correction for dispersion
wvrflag = [''] # Flag the WVR data for these antenna(s) as bad and replace
# its data with interpolated values
statfield = '' # Compute the statistics (Phase RMS, Disc) on this field only
statsource = '' # Compute the statistics (Phase RMS, Disc) on this source only
smooth = '' # Smooth calibration solution on the given timescale
scale = 1.0 # Scale the entire phase correction by this factor
spw = [] # List of the spectral window IDs for which solutions should
# be saved into the caltable
wvrspw = [] # List of the spectral window IDs from which the WVR data
# should be taken
reversespw = '' # Reverse the sign of the correction for the listed SPWs
# (only needed for early ALMA data before Cycle 0)
cont = False # Estimate the continuum (e.g., due to clouds) (experimental)
maxdistm = 500.0 # maximum distance (m) of an antenna used for interpolation
# for a flagged antenna
minnumants = 2 # minimum number of near antennas (up to 3) required for
# interpolation
mingoodfrac = 0.8 # If the fraction of unflagged data for an antenna is below
# this value (0. to 1.), the antenna is flagged.
usefieldtab = False # derive the antenna AZ/EL values from the FIELD rather than
# the POINTING table

```

The task `wvrgcal` generates a gain table based on Water Vapor Radiometer (WVR) data and is used for ALMA data reduction. It is an interface to the executable “wvrgcal” which is part of the CASA 4.3 distribution and can also be called from outside CASA. The wvrgcal software is based on the `libair` and `libnmin` libraries which were developed by Bojan Nikolic at the University of Cambridge as part of EU FP6 ALMA Enhancement program.

CASA 4.3 contains version 2.0 of wvrgcal. Source code of the stand-alone package and links to documentation up to version 1.2.1 can be found online¹. In particular, there are three ALMA memos (number 587, 588, and 593 (submitted)) which describe the algorithms implemented in the software. They can be found at at the ALMA Memo Series².

With wvrgcal version 2.0, maintenance of the tool has now fullt transitioned to ESO. Newly added features include full support for flags, and the new parameters `spw`, `wvrspw`, `maxdistm`, `mingoodfrac`, and `usefieldtab`.

Briefly, wvrgcal follows a Bayesian approach to calculate the coefficients that convert the outputs of the ALMA 183 GHz water-vapor radiometers (mounted on each antenna) into estimates of path fluctuations which can then be used to correct the observed interferometric visibilities.

The CASA task interface to wvrgcal follows closely the interface of the shell executable at the same time staying within the CASA task parameter conventions.

In ALMA data, the WVR measurements belonging to a given observation are contained in the ASDM for that observation. After conversion to an MS using `importasdm`, the WVR information

¹<http://www.mrao.cam.ac.uk/~bn204/alma/wvrsoft.html>

²<http://www.alma.cl/almamemos>

can be found in separate spectral windows (as of November 2014, it is still spectral window id 0 only). This spectral window must be present in the MS for `wvrgcal` to work.

The various features of `wvrgcal` are then controlled by a number of task parameters (see the list above). They have default values which will work for ALMA data. An example for a typical `wvrgcal` call can be found in the ALMA CASA guide for the NGC 3256 analysis:

```
wvrgcal(vis='uid__A002_X1d54a1_X5.ms', caltable='cal-wvr-uid__A002_X1d54a1_X5.W',
        toffset=-1, segsource=True, tie=["Titan,1037-295,NGC3256"], statsource="1037-295")
```

Here, `vis` is the name of input visibility file (which as mentioned above also contains the WVR data in spectral window 0) and `caltable` is the name of the output gain calibration table.

`toffset` is the known time offset in seconds between the WVR measurements and the visibility integrations they are valid for. For ALMA, this offset is presently -1 s (which is also the default value).

The parameter `segsource` (segregate source) controls whether separate coefficients are calculated for each source. The default value True is the recommended one for ALMA. When `segsource` is True, the subparameter `tie` is available. It permits to form groups of sources for which common coefficients are calculated as well as possible. The `tie` parameter ensures best possible phase transfer between a group of sources. In general it is recommended to tie together all of the sources in a single Science Goal (in ALMA speak) and their phase calibrator(s). The recommended maximum angular distance up to which two sources can be tied is 15°.

The parameter `statsource` controls for which sources statistics are calculated and displayed in the logger. This has no influence on the generated calibration table.

`wvrgcal` respects the flags in the Main and ANTENNA table of the MS. The parameter `mingoodfrac` lets the user set a requirement on the minimum fraction of good measurements for accepting the WVR data from an antenna. If antennas are flagged, their WVR solution is interpolated from the three nearest neighbouring antennas. This process can be controlled with the new parameters `maxdistm` and `minnumants`. The former sets the maximum distance an antenna used for interpolation may have from the flagged one. And `minnumants` sets how many near antennas there have to be for interpolation to take place.

For more details on the WVR Phase correction, see also the the ALMA Memo “Quality Control of WVR Phase Correction Based on Differences Between WVR Channels” by B. Nikolic, R. C. Bolton & J. S. Richer³, see also ALMA memo #593⁴.

4.3.8.1 Statistical parameters shown in the logger output of `wvrgcal`

`wvrgcal` writes out a variety of information to the logger, including various statistical measures of the performance. This allows the user to judge whether WVR correction is appropriate for the

³<http://www.casa.nrao.edu/Memos/memoqachannels.pdf>

⁴https://science.nrao.edu/facilities/alma/aboutALMA/Technology/ALMA_Memo_Series/alma593/abs593

ms, to check whether any antennas have problematic WVR values, and to examine the predicted performance of the WVR correction when applied.

For each set of correction coefficients which are calculated (the number of coefficient sets are controlled by the parameters `nsol`, `segsources` and `tie`), the `wvrgcal` output to the logger first of all shows the time sample, the individual temperatures of each of the four WVR channels and the elevation of the source in question at that time.

For each of these coefficient sets, it then gives the evidence of the bayesian parameter estimation, the calculated precipitable water vapour (PWV) and its error in mm, and the correction coefficients found for each WVR channel (`dTdL`).

The output then shows the statistical information about the observation. First of all it gives the start and end times for the parts of the observation used to calculate these statistics (controlled by `segsources`). It then shows a break down for each of the antennas in the data set. This gives the antenna name and number; whether or not it has a WVR (column `WVR`); whether or not it has been flagged (column `FLAG`); the RMS of the path length variation with time towards that antenna (column `RMS`); and the discrepancy between the RMS path length calculated separately for different WVR channels (column `Disc.`). These values allow the user to see if an individual WVR appears to have been suffering from problems during the observation, and to flag that antenna using `wvrflag` if necessary.

This discrepancy value, `Disc.`, can in addition be used as a simple diagnostic tool to evaluate whether or not the WVR correction caltable created by `wvrgcal` should be applied. In the event of the WVR observations being contaminated by strong cloud emission in the atmosphere, the attempt by `wvrgcal` to fit the water vapour line may not be successful, and applying the produced calibration table can in extreme cases reduce the quality of the data. However, these weather conditions should be identified by a high value in the discrepancy column produced when running `wvrgcal`.

Although there have not currently been enough cases checked to give definitive limits, the available data sets as of summer 2012 suggest that discrepancy values of greater than a 1000 microns usually indicate strong cloud contamination of the WVR data, and the output calibration table should probably not be applied. If the values are between 100 and 1000 microns, then the user should manually examine the phases before and after applying the caltable to decide if WVR correction is appropriate.

After the antenna-by-antenna statistics, the output then displays some estimates of the performance of the `wvrgcal` correction. These are the thermal contribution from the water vapour to the path fluctuations per antenna (in microns), the largest path fluctuation found on a baseline (in microns), and the expected error on the path length calculated for each baseline due to the error in the coefficients (in microns).

4.3.8.2 Antenna position calculation

The information about antenna pointing direction is by default taken from the `POINTING` table. Should this table not be present for some reason, the user can instead switch to determining the

antenna positions from the phase directions in the FIELD table (under the assumption that all antennas were pointing ideally). The switch is performed by setting the parameter `usefieldtab` to `True`.

4.3.8.3 Spectral window selection

By default `wvrgcal` has traditionally put solutions for *all* spectral windows of the MS into the output calibration table. Since usually only the spectral windows are of interest in which the science target and the calibrators were observed, it is not necessary to store solutions for other spectral windows.

The spectral windows for which solutions are stored can be selected with the parameter `spw`, e.g. `spw = [17,19,21,23]` will make `wvrgcal` write only solutions for spectral windows 17, 19, 21, and 23.

Should ALMA in the future use other spectral windows than id 0 to store the WVR information, then the input WVR spectral window can be selected using the parameter `wvrspw`. The syntax is the same as for parameter `spw`.

4.3.9 Ionospheric corrections

CASA 4.3 introduces initial support for on-axis ionospheric corrections, using time- and direction-dependent total electron content (TEC) information obtained from the internet. The correction includes the dispersive delay ($\propto \nu^{-1}$) delay and Faraday rotation ($\propto \nu^{-2}$) terms. These corrections are most relevant at observing frequencies less than ~ 5 GHz. When relevant, the ionosphere correction table should be generated at the beginning of a reduction along with other calibration priors (antenna position errors, gain curve, opacity, etc.), and carried through all subsequent calibration steps. Formally, the idea is that the ionospheric effects (as a function of time and on-axis direction) will be nominally accounted for by this calibration table, and thus not spuriously leak into gain and bandpass solves, etc. In practice, the quality of the ionospheric correction is limited by the relatively sparse sampling (in time and direction) of the available TEC information. Especially active ionospheric conditions may not be corrected very well. Also, direction-dependent ionosphere corrections are not yet supported. (Various improvements are under study for future releases.)

To generate the ionosphere correction table, first import a helper function from the `casapy` recipes repository:

```
from recipes import tec_maps
```

Then, generate a TEC surface image:

```
tec_maps.create(vis='mydata.ms',doplot=T,imname='iono')
```

This function goes to the web to obtain TEC information for the observing date and location, and generates a time-dependent CASA image containing this information. The string specified for `imname` is used as a prefix for two output images, with suffixes `.IGS_TEC.im` (the actual TEC image) and `.IGS_RMS_TEC.im` (a TEC error image). If `imname` is unspecified, the MS name (from `vis`) will be used as the prefix.

The quality of the retrieved TEC information improves with time after the observing date, becoming optimal 1-2 weeks later. Both images can be viewed as a movie in the CASA viewer. If `doplot=T`, the function will also produce a plot of the TEC as a function of time in a vertical direction over the observatory.

Finally, to generate the ionosphere correction caltable, pass the `.IGS_TEC.im` image into `gencal`, using `caltype='tecim'`:

```
gencal(vis='mydata.ms', caltable='tec.cal', caltype='tecim', infile='iono.IGS_TEC.im')
```

This iterates through the dataset and samples the zenith angle-dependent projected line-of-sight TEC for all times in the observation, storing the result in a standard CASA caltable. Plotting this caltable will show how the TEC varies between observing directions for different fields and times, in particular how it changes as zenith angle changes, and including the nominal difference between science targets and calibrators.

This caltable should then be used as a prior in all subsequent calibration solves, and included in the final `applycal`.

A few warnings:

- The TEC information obtained from the web is relatively poorly sampled in time and direction, and so will not always describe the details of the ionospheric corruption, especially during active periods.
- For instrumental polarization calibration, it is recommended that an *unpolarized* calibrator be used; polarized calibrators may not yield as accurate a solution since the ionospheric corrections are not yet used properly in the source polarization portion of the solve.

Special thanks are due to Jason Kooi (UIowa) for his contributions to ionospheric corrections in CASA.

4.3.10 Other *a priori* Calibrations and Corrections

Other *a priori* calibrations will be added to the `calibrator` (`cb`) tool in the near future. These will include instrumental line-length corrections, etc. Where appropriate, solving capabilities for these effects will also be added.

4.4 Solving for Calibration — Bandpass, Gain, Polarization

The `gaincal`, `bandpass`, `polcal`, and `blcal` tasks actually solve for the unknown calibration parameters from the visibility data obtained on calibrator sources, placing the results in a calibration table. They take as input an MS, and a number of parameters that specify any prior calibration or previous calibration tables to pre-apply before computing the solution. These are placed in the proper sequence of the Measurement Equation automatically.

We first discuss the parameters that are in common between many of the calibration tasks. Then we describe each solver in turn.

4.4.1 Common Calibration Solver Parameters

There are a number of parameters that are in common between the calibration “solver” tasks. These also appear in some of the other calibration manipulation and application tasks.

4.4.1.1 Parameters for Specification : `vis` and `caltable`

The input Measurement Set and output table are controlled by the following parameters:

```
vis          =      '' # Name of input visibility file
caltable    =      '' # Name of output calibration table
```

The MS name is input in `vis`. If it is highlighted red in the inputs (§ 1.4.5.4) then it does not exist, and the task will not execute. Check the name and path in this case.

The output table name is placed in `caltable`. Be sure to give a unique name to the output table, or be careful. If the table exists, then what happens next will depend on the task and the values of other parameters (e.g. § 4.4.1.6). The task may not execute giving a warning that the table already exists, or will go ahead and overwrite the solutions in that table, or append them. Be careful.

4.4.1.2 Selection: `field`, `spw`, `selectdata`, `intent`, and `observation`

Selection is controlled by the parameters:

```
field       =      '' # field names or index of calibrators: ''==>all
spw        =      '' # spectral window:channels: ''==>all
intent     =      '' # Select observing intent
selectdata =      False # Other data selection parameters
```

Field and spectral window selection are so often used, that we have made these standard parameters `field` and `spw` respectively. `intent` is the scan intent that was specified when the observations were set up. They typically describe what was intended with a specific scan, i.e. a flux or phase calibration, a bandpass, a pointing, an observation of your target, or something else or a combination.

The format for the scan intents of your observations are listed in the logger when you run `listobs`. Minimum matching with wildcards will work, like `'*BANDPASS*`'. This is especially useful when multiple intents are attached to scans. Finally, `observation` is an identifier to distinguish between different observing runs, mainly used for ALMA.

The `selectdata` parameter expands as usual, uncovering other selection sub-parameters:

```
selectdata      =      True      # data selection parameters
  field         =      ''        # field names or field index
                                     # numbers (blank for all)
  spw           =      ''        # spectral windows:channels (blank for all)
  timerange     =      ''        # time range (blank for all)
  uvrange       =      ''        # uv range (blank for all)
  antenna       =      ''        # antenna/baselines (blank for all)
  scan          =      ''        # scan numbers (blank for all)
  correlation   =      ''        # correlations (blank for all)
  array         =      ''        # (sub)array numbers (blank for all)
  observation   =      ''        # Select by observation ID(s)
  msselect     =      ''        # MS selection (blank for all)
```

Note that if `selectdata=False` these parameters are not used when the task is executed, even if set underneath.

The most common `selectdata` parameter to use is `uvrange`, which can be used to exclude longer baselines if the calibrator is resolved, or short baselines of the calibrator contains extended flux not accounted for in the model (e.g. § 4.3.5.1).

See § 2.3 for more on the selection parameters.

4.4.1.3 Prior Calibration and Correction: `parang`

These parameters control the on-the-fly application of various calibration or effect-based corrections prior to the solving process.

The `parang` parameter turns on the application of the antenna-based parallactic angle correction ('P') in the measurement equation. This is necessary for polarization calibration and imaging, or for cases where the parallactic angles are different for geographically spaced antennas and it is desired that the ordinary gain calibration not absorb the inter-antenna parallactic angle phase. When dealing with only the parallel-hand data (e.g. RR, LL, XX, YY), and an unpolarized calibrator model for a co-located array (e.g. the VLA or ALMA), you can set `parang=False` and save some computational effort. Otherwise, set `parang=True` to apply this correction.

See § 4.3 for more on **Prior Calibration**, including how to invoke `gaincurve` and `opacity` correction using `gencal`.

4.4.1.4 Previous Calibration: `gaintable`, `gainfield`, `interp` and `spwmap`

Calibration tables that have already been determined can also be applied before solving for the new table:


```

docallib      = False      # Use traditional cal apply parameters
  gaintable   = []         # Gain calibration table(s) to apply on the fly
  gainfield   = []         # Select a subset of calibrators from gaintable(s)
  interp      = []         # Interpolation mode (in time) to use for each gaintable
  spwmap      = []         # Spectral windows combinations to form for gaintable(s)

```

This is controlled by the `gaintable` parameter, which takes a string or list of strings giving one or more calibration tables to pre-apply. For example,

```
gaintable = ['ngc5921.bcal', 'ngc5921.gcal']
```

specifies two tables, in this case bandpass and gain calibration tables respectively.

The other parameters key off `gaintable`, taking single values or lists, with an entry for each table in `gaintable`. The order is given by that in `gaintable`.

The `gainfield` parameter specifies which fields from the respective `gaintable` to select for apply. This is a list, with each entry a string or list of strings. The default `''` for an entry means to use all in that table. For example,

```
gaintable = ['ngc5921.bcal', 'ngc5921.gcal']
gainfield = [ '1331+305', ['1331+305', '1445+099'] ]
```

or using indices

```
gainfield = [ '0', ['0', '1'] ]
```

to specify the field `'1331+305'` from the table `'ngc5921.bcal'` and fields `'1331+305'` and `'1445+099'` from the second table `'ngc5921.gcal'`. We could also have wildcarded the selection, e.g.

```
gainfield = [ '0', '*' ]
```

taking all fields from the second table. And of course we could have used the default

```
gainfield = [ '0', '' ]
```

or even

```
gainfield = [ '0' ]
```

which is to take all for the second table in `gaintable`. In addition, `gainfield` can be specified by

```
gainfield = [ 'nearest' ]
```

which selects the calibrator that is the spatially closest (in sky coordinates) to each of the selected MS fields specified in the `field` parameter. This is particularly useful for running `applycal` with a number of different sources to be calibrated in a single run.

The `interp` parameter chooses the interpolation scheme to be used when pre-applying the solution in the tables. Interpolation in both time and frequency (for channel-dependent calibrations) are supported. The choices are currently `'nearest'` and `'linear'`, and `'nearest'`, `'linear'`, `cubic`, and `spline` for frequency-dependent interpolation. Frequency-dependent interpolation is only relevant for channel-dependent calibration tables (like bandpasses) that are undersampled in frequency relative to the data.

- `'nearest'` just picks the entry nearest in time or freq to the visibility in question;
- `'linear'` interpolation calibrates each datum with calibration phases and amplitudes linearly interpolated from neighboring values. In the case of phase, this mode will assume that phase never jumps more than 180° between neighboring points, and so undersampled cycle-slips will not be corrected for. Solutions will not be *extrapolated* arbitrarily in time or frequency for data before the first solution or after the last solution; such data will be calibrated using `'nearest'` to avoid unreasonable extrapolations.
- `'cubic'` interpolation forms a 3rd-order polynomial that passes through the nearest 4 calibration samples (separately in phase and amplitude)
- `'spline'` interpolation forms a cubic spline that passes through the nearest 4 calibration samples (separately in phase and amplitude)

For each gaintable, specify the interpolation style in quotes, with the frequency-dependent interpolation style specified after a comma, if relevant.

If the uncalibrated phase is changing rapidly, a `'nearest'` interpolation is not desirable. Usually, `interp='linear'` is the best choice. For example,

```
gaintable=['gain','bandpass']
interp = [ 'nearest', 'linear,spline' ]
```

uses nearest “interpolation” on the first table, and linear (in time) and spline (in freq) on the second.

The `spwmap` parameter sets the spectral window combinations to form for the `gaintable(s)`. This is a list, or a list of lists, of integers giving the `spw` IDs to map. There is one list for each table in `gaintable`, with an entry for each ID in the MS. For example,

```
spwmap=[0,0,1,1] # apply from spw=0 to 0,1 and 1 to 2,3
```

for an MS with `spw=0,1,2,3`. For multiple `gaintable`, use lists of lists, e.g.

```
spwmap=[ [0,0,1,1], [0,1,0,1] ] # 2nd table spw=0 to 0,2 and 1 to 1,3
```

4.4.1.5 Solving: solint, combine, preavg, refant, minblperant, minsnr

The parameters controlling common aspects of the solution are:

```

solint      =      'inf'      # Solution interval: egs. 'inf', '60s' (see help)
combine     =      'scan'     # Data axes which to combine for solve (obs, scan,
                               # spw, and/or field)
preavg      =      -1.0      # Pre-averaging interval (sec) (rarely needed)
refant      =      ''        # Reference antenna name(s)
minblperant =      4         # Minimum baselines _per antenna_ required for solve
minsnr      =      3.0      # Reject solutions below this SNR

```

The time and frequency (if relevant) solution interval is given by `solint`. Optionally a frequency interval for each solution can be added after a comma, e.g. `solint='60s,300Hz'`. Time units are in seconds unless specified differently. Frequency units can be either channels or Hz and only make sense for bandpass of frequency dependent polarization calibration. The special values `'inf'` and `-1` specify an “infinite” solution interval encompassing the entire dataset, while `'int'` or zero specify a solution every integration. You can use time quanta in the string, e.g. `solint='1min'` and `solint='60s'` both specify solution intervals of one minute. Note that `'m'` is a unit of distance (meters); `'min'` must be used to specify minutes. The `solint` parameter interacts with `combine` to determine whether the solutions cross scan or field boundaries.

The parameter controlling the scope of the solution is `combine`. For the default `combine=''` solutions will break at `obsId`, `scan`, `field`, and `spw` boundaries. Specification of any of these in `combine` will extend the solutions over the boundaries (up to the `solint`). For example, `combine='spw'` will combine spectral windows together for solving, while `combine='scan'` will cross scans, and `combine='obs,scan'` will use data across different observation IDs and scans (usually, `obsIds` consist of many scans, so it is not meaningful to combine `obsIds` without also combining scans). Thus, to do scan-based solutions (single solution for each scan), set

```

solint = 'inf'
combine = ''

```

while

```

solint = 'inf'
combine = 'scan'

```

will make a single solution for the entire dataset (for a given field and spw).

```

solint = 'inf,30ch'

```

will calculate a bandpass solution for each scan, averaging over 30 channels.

You can specify multiple choices for combination:

```

combine = 'scan,spw'

```

for example.

Alert: Care should be exercised when using `combine='spw'` in cases where multiple groups of concurrent spectral windows are observed as a function of time. Currently, only *one* aggregate spectral window can be generated in a single calibration solve execution, and the meta-information for this spectral window is calculated from all selected MS spectral windows. To avoid incorrect calibration meta-information, each spectral window group should be calibrated independently (also without using `append=True`). Additional flexibility in this area will be supported in a future version.

The reference antenna is specified by the `refant` parameter. A list of antennas can be provided to this parameter and if the first antenna is not present in the data, the next antenna in the list will be used, etc. It is useful to “lock” the solutions with time, effectively rotating (after solving) the phase of the gain solutions for all antennas such that the reference antenna’s phase is constant at zero. If the selected antenna drops out, another antenna will be selected for ongoing consistency in time (at its “current” value) until the `refant` returns, usually at a new value (not zero), which will be kept fixed thenceforth. You can also run without a reference antenna, but in this case the solutions will formally float with time; in practice, the first antenna will be approximately constant near zero phase. It is usually prudent to select an antenna in the center of the array that is known to be particularly stable, as any gain jumps or wanders in the `refant` will be transferred to the other antenna solutions. Also, it is best to choose a reference antenna that never drops out.

Setting a `preavg` time (only needed in `polcal`) will let you average data over periods shorter than the solution interval first before solving on longer timescales.

The minimum signal-to-noise ratio allowed for an acceptable solution is specified in the `minsnr` parameter. Default is `minsnr=3`. The `minblperant` parameter sets the minimum number of baselines to other antennas that must be preset for each antenna to be included in a solution. This enables control of the constraints that a solution will require for each antenna.

4.4.1.6 Action: `append` and `solnorm`

The following parameters control some things that happen after solutions are obtained:

```
solnorm      =      False  #   Normalize solution amplitudes post-solve.
append       =      False  #   Append solutions to (existing) table.
                                     #   False will overwrite.
```

The `solnorm` parameter toggles on the option to normalize the solution after the solutions are obtained. The exact effect of this depends upon the type of solution. Not all tasks use this parameter.

One should be aware when using `solnorm` that if this is done in the last stage of a chain of calibration, then the part of the calibration that is “normalized” away will be lost. It is best to use this in early stages (for example in a first bandpass calibration) so that later stages (such as final gain calibration) can absorb the lost normalization scaling. It is not strictly necessary to use `solnorm=True` at all, but is sometimes helpful if you want to have a normalized bandpass for example.

The `append` parameter, if set to `True`, will append the solutions from this run to existing solutions in `caltable`. Of course, this only matters if the table already exists. If `append=False` and `caltable` exists, it will overwrite.

The `append` parameter should be used with care, especially when also using `combine` in non-trivial ways. E.g., calibration solves will currently (CASA 4.5) refuse to append incongruent aggregate spectral windows (e.g., observations with more than one group of concurrent spectral windows). This limitation arises from difficulty determining the appropriate spectral window fan-out on apply, and will be relaxed in a future version.

4.4.2 Spectral Bandpass Calibration (bandpass)

For channelized data, it is usually desirable to solve for the gain variations in frequency as well as in time. Variation in frequency arises as a result of non-uniform filter passbands or other frequency-dependent effects in signal transmission. It is usually the case that these frequency-dependent effects vary on timescales much longer than the time-dependent effects handled by the gain types 'G' and 'T'. Thus, it makes sense to solve for them as a separate term: 'B', using the `bandpass` task.

The inputs to `bandpass` are:

```
# bandpass :: Calculates a bandpass calibration solution
vis                =      ''          # Name of input visibility file
caltable           =      ''          # Name of output gain calibration table
field              =      ''          # Select field using field id(s) or field name(s)
spw                =      ''          # Select spectral window/channels
intent             =      ''          # Select observing intent
selectdata         =      True        # Other data selection parameters
    timerange      =      ''          # Select data based on time range
    uvrange        =      ''          # Select data within uvrange (default units meters)
    antenna        =      ''          # Select data based on antenna/baseline
    scan           =      ''          # Scan number range
    observation    =      ''          # Select by observation ID(s)
    msselect       =      ''          # Optional complex data selection (ignore for now)

solint             =      'inf'       # Solution interval in time[,freq]
combine            =      'scan'      # Data axes which to combine
                                     # for solve (obs, scan, spw, and/or field)
refant             =      ''          # Reference antenna name(s)
minblperant       =      4           # Minimum baselines _per
                                     # antenna_ required for solve
minsnr             =      3.0        # Reject solutions below this
                                     # SNR (only applies for bandtype = B)
solnorm            =      False       # Normalize average solution amplitudes to 1.0
bandtype           =      'B'        # Type of bandpass solution (B or BPOLY)
    fillgaps       =      0           # Fill flagged solution channels by interpolation

smodel             =      []          # Point source Stokes parameters for source model.
append             =      False       # Append solutions to the (existing) table
```

```

docallib      =      False      # Use callib or traditional cal apply parameters
  gainable    =      []         # Gain calibration table(s) to apply on the fly
  gainfield   =      []         # Select a subset of calibrators from gainable(s)
  interp      =      []         # Interpolation mode (in
                                #   time) to use for each gainable
  spwmap      =      []         # Spectral windows
                                #   combinations to form for gainable(s)

parang        =      False      # Apply parallactic angle correction

```

Many of these parameters are in common with the other calibration tasks and are described above in § 4.4.1.

The `bandtype` parameter selects the type of solution used for the bandpass. The choices are 'B' and 'BPOLY'. The former solves for a complex gain in each channel in the selected part of the MS. See § 4.4.2.2 for more on 'B'. The latter uses a polynomial as a function of channel to fit the bandpass, and expands further to reveal a number of sub-parameters See § 4.4.2.3 for more on 'BPOLY'.

It is usually best to solve for the bandpass in channel data before solving for the gain as a function of time. However, if the gains of the bandpass calibrator observations are fluctuating over the timerange of those observations, then it can be helpful to first solve for the gains of that source with `gaincal`, and input these to `bandpass` via `gainable`. See more below on this strategy.

We now describe the issue of bandpass normalization, followed by a description of the options `bandtype='B'` and `bandtype='BPOLY'`.

4.4.2.1 Bandpass Normalization

The `solnorm` parameter (§ 4.4.1.6) deserves more explanation in the context of the bandpass. Most users are used to seeing a normalized bandpass, where the mean amplitude is unity and fiducial phase is zero. The toggle `solnorm=True` allows this. However, the parts of the bandpass solution normalized away will be still left in the data, and thus you should not use `solnorm=True` if the `bandpass` calibration is the end of your calibration sequence (e.g. you have already done all the gain calibration you want to). Note that setting `solnorm=True` will NOT rescale any previous calibration tables that the user may have supplied in `gainable`.

You can safely use `solnorm=True` if you do the bandpass first (perhaps after a throw-away initial gain calibration) as we suggest above in § 4.2, as later gain calibration stages will deal with this remaining calibration term. This does have the benefit of isolating the overall (channel independent) gains to the following `gaincal` stage. It is also recommended for the case where you have multiple scans on possibly different bandpass calibrators. It may also be preferred when applying the bandpass before doing `gaincal` and then `fluxscale` (§ 4.4.4), as significant variation of bandpass among antennas could otherwise enter the gain solution and make (probably subtle) adjustments to the flux scale.

We finally note that `solnorm=False` at the bandpass step in the calibration chain will still in the end produce the correct results. It only means that there will be a part of what we usually think of

the gain calibration inside the bandpass solution, particularly if `bandpass` is run as the first step.

4.4.2.2 B solutions

Calibration type 'B' differs from 'G' only in that it is determined for each channel in each spectral window. It is possible to solve for it as a function of time, but it is most efficient to keep the 'B' solving timescale as long as possible, and use 'G' or 'T' for frequency-independent rapid time-scale variations.

The 'B' solutions are limited by the signal-to-noise ratio available per channel, which may be quite small. It is therefore important that the data be coherent over the time-range of the 'B' solutions. As a result, 'B' solutions are almost always preceded by an initial 'G' or 'T' solve using `gaincal` (§ 4.4.3). In turn, if the 'B' solution improves the frequency domain coherence significantly, a 'G' or 'T' solution following it will be better than the original.

For example, to solve for a 'B' bandpass using a single short scan on the calibrator, then

```
default('bandpass')

vis = 'n5921.ms'
caltable = 'n5921.bcal'
gaintable = ''           # No gain tables yet
gainfield = ''
interp = ''
field = '0'             # Calibrator 1331+305 = 3C286 (FIELD_ID 0)
spw = ''               # all channels
selectdata = False     # No other selection
bandtype = 'B'         # standard time-binned B (rather than BPOLY)
solint = 'inf'         # set solution interval arbitrarily long
refant = '15'          # ref antenna 15 (=VLA:N2) (ID 14)

bandpass()
```

On the other hand, we might have a number of scans on the bandpass calibrator spread over time, but we want a single bandpass solution. In this case, we could solve for and then pre-apply an initial gain calibration, and let the bandpass solution cross scans:

```
gaintable = 'n5921.init.gcal' # Our previously determined G table
gainfield = '0'
interp = 'linear'           # Do linear interpolation
solint = 'inf'              # One interval over dataset
combine = 'scan'           # Solution crosses scans
```

Note that we obtained a bandpass solution for all channels in the MS. If explicit channel selection is desired, for example some channels are useless and can be avoided entirely (e.g. edge channels or those dominated by Gibbs ringing), then `spw` can be set to select only these channels, e.g.

```
spw = '0:4~59'             # channels 4-59 of spw 0
```

This is not so critical for 'B' solutions as for 'BPOLY', as each channel is solved for independently, and poor solutions at edges can be ignored.

If you have multiple time solutions, then these will be applied using whatever time interpolation scheme is specified in later tasks.

The `combine` parameter (§ 4.4.1.5) can be used to combine data across spectral windows, scans, and fields.

4.4.2.3 BPOLY solutions

For some observations, it may be the case that the SNR per channel is insufficient to obtain a usable per-channel 'B' solution. In this case it is desirable to solve instead for a best-fit functional form for each antenna using the `bandtype='BPOLY'` solver. The 'BPOLY' solver naturally enough fits (Chebychev) polynomials to the amplitude and phase of the calibrator visibilities as a function of frequency. Unlike ordinary 'B', a single common 'BPOLY' solution will be determined for all spectral windows specified (or implicit) in the selection. As such, it is usually most meaningful to select individual spectral windows for 'BPOLY' solves, unless groups of adjacent spectral windows are known *a priori* to share a single continuous bandpass response over their combined frequency range (e.g., PdBI data).

The 'BPOLY' solver requires a number of unique sub-parameters:

```
bandtype      =      'BPOLY'  #   Type of bandpass solution (B or BPOLY)
  degamp      =           3  #   Polynomial degree for BPOLY amplitude solution
  degphase    =           3  #   Polynomial degree for BPOLY phase solution
  visnorm     =      False  #   Normalize data prior to BPOLY solution
  maskcenter  =           0  #   Number of channels in BPOLY to avoid in center of band
  maskedge    =           0  #   Percent of channels in BPOLY to avoid at each band edge
```

The `degamp` and `degphase` parameters indicate the polynomial degree desired for the amplitude and phase solutions. The `maskcenter` parameter is used to indicate the number of channels in the center of the band to avoid passing to the solution (e.g., to avoid Gibbs ringing in central channels for PdBI data). The `maskedge` drops beginning and end channels. The `visnorm` parameter turns on normalization before the solution is obtained (rather than after for `solnorm`).

The `combine` parameter (§ 4.4.1.5) can be used to combine data across spectral windows, scans, and fields.

Note that `bandpass` will allow you to use multiple `fields`, and can determine a single solution for all specified fields using `combine='field'`. If you want to use more than one field in the solution it is prudent to use an initial `gaincal` using proper flux densities for all sources (not just 1Jy) and use this table as an input to `bandpass` because in general the phase towards two (widely separated) sources will not be sufficiently similar to combine them, and you want the same amplitude scale. If you do not include amplitude in the initial `gaincal`, you probably want to set `visnorm=True` also to take out the amplitude normalization change. Note also in the case of multiple `fields`, that the 'BPOLY' solution will be labeled with the field ID of the first `field` used in the 'BPOLY'

solution, so if for example you point `plotcal` at the name or ID of one of the other fields used in the solution, `plotcal` does not plot.

For example, to solve for a 'BPOLY' (5th order in amplitude, 7th order in phase), using data from field 2, with G corrections pre-applied:

```
bandpass(vis='data.ms',          # input data set
         caltable='cal.BPOLY',  #
         spw='0:2~56',         # Use channels 3-57 (avoid end channels)
         field='0',            # Select bandpass calibrator (field 0)
         bandtype='BPOLY',     # Select bandpass polynomials
         degamp=5,             # 5th order amp
         degphase=7,           # 7th order phase
         gaintable='cal.G',    # Pre-apply gain solutions derived previously
         refant='14')          #
```

4.4.2.4 What if the bandpass calibrator has a significant slope?

The bandpass calibrator can have a spectral slope that will change the spectral properties of the solutions. If the slope is significant, the best way is to model the slope and store that model in the bandpass calibrator MS. To do so, go through the normal steps of `bandpass` and the `gaincal` runs on the bandpass and flux calibrators, followed by `setjy` of the flux calibrator. The next step would be to use `fluxscale` on the bandpass calibrator to derive the slope of it. `fluxscale` can store this information in a python dictionary which is subsequently fed into a second `setjy` run, this time using the bandpass calibrator as the source and the derived slope (the python dictionary) as input. This step will create a source model with the correct overall spectral slope for the bandpass. Finally, rerun `bandpass` and all other calibration steps again, making use of the newly created internal bandpass model.

4.4.3 Complex Gain Calibration (`gaincal`)

The fundamental calibration to be done on your interferometer data is to calibrate the antenna-based gains as a function of time. Some of these calibrations are known beforehand (“a priori”) and others must be determined from observations of calibrators, or from observations of the target itself (“self-calibration”).

It is best to have determined a (constant or slowly-varying) “bandpass” from the frequency channels by solving for the bandpass (see above). Thus, the `bandpass` calibration table would be input to `gaincal` via the `gaintable` parameter (see below).

The `gaincal` task has the following inputs:

```
# gaincal :: Determine temporal gains from calibrator observations
vis          =      ''          # Name of input visibility file
caltable     =      ''          # Name of output gain calibration table
field       =      ''          # Select field using field id(s) or field name(s)
spw         =      ''          # Select spectral window/channels
```

```

intent          =      ''          # Select observing intent
selectdata     =      True        # Other data selection parameters
  timerange    =      ''          # Select data based on time range
  uvrange      =      ''          # Select data within uvrange (default units meters)
  antenna      =      ''          # Select data based on antenna/baseline
  scan         =      ''          # Scan number range
  observation  =      ''          # Select by observation ID(s)
  msselect     =      ''          # Optional complex data selection (ignore for now)

solint         =      'inf'       # Solution interval: egs. 'inf', '60s' (see help)
combine       =      ''          # Data axes which to combine
  # for solve (obs, scan, spw, and/or field)
preavg        =      -1.0        # Pre-averaging interval (sec) (rarely needed)
refant        =      ''          # Reference antenna name(s)
minblperant   =      4           # Minimum baselines _per antenna_ required for solve
minsnr        =      3.0         # Reject solutions below this SNR
solnorm       =      False       # Normalize average solution
  # amplitudes to 1.0 (G, T only)
gaintype      =      'G'         # Type of gain solution (G,T,GSPLINE,K,KCROSS)
smodel        =      []          # Point source Stokes parameters for source model.
calmode       =      'ap'        # Type of solution: ('ap', 'p', 'a')
append        =      False       # Append solutions to the (existing) table
docallib      =      False       # Use callib or traditional cal apply parameters
  gaintable    =      []          # Gain calibration table(s) to apply on the fly
  gainfield    =      []          # Select a subset of calibrators from gaintable(s)
  interp       =      []          # Temporal interpolation for
  # each gaintable (=linear)
  spwmap       =      []          # Spectral windows
  # combinations to form for gaintable(s)

parang        =      False       # Apply parallactic angle correction on the fly

```

Data selection is done through the standard `field`, `spw`, `intent`, and `selectdata` expandable sub-parameters (see § 2.3). The bulk of the other parameters are the standard solver parameters. See § 4.4.1 above for a description of these.

The `gaintype` parameter selects the type of gain solution to compute. The choices are `'T'`, `'G'`, and `'GSPLINE'`. The `'G'` and `'T'` options solve for independent complex gains in each solution interval (classic AIPS style), with `'T'` enforcing a single polarization-independent gain for each co-polar correlation (e.g. RR and LL, or XX and YY) and `'G'` having independent gains for these. See § 4.4.3.1 for a more detailed description of `'G'` solutions, and § 4.4.3.2 for more on `'T'`. The `'GSPLINE'` fits cubic splines to the gain as a function of time. See § 4.4.3.3 for more on this option.

4.4.3.1 Polarization-dependent Gain (G)

Systematic time-dependent complex gain errors are almost always the dominant calibration effect, and a solution for them is almost always necessary before proceeding with any other calibration. Traditionally, this calibration type has been a catch-all for a variety of similar effects, including: the

relative amplitude and phase gain for each antenna, phase and amplitude drifts in the electronics of each antenna, amplitude response as a function of elevation (gain curve), and tropospheric amplitude and phase effects. In CASA, it is possible to handle many of these effects separately, as available information and circumstances warrant, but it is still possible to solve for the net effect using calibration type G.

Generally speaking, type G can represent any per-spectral window multiplicative polarization- and time-dependent complex gain effect downstream of the polarizers. (Polarization- and time-*independent* effects *upstream* of the polarizers may also be treated implicitly with G.) Multi-channel data (per spectral window) will be averaged in frequency before solving (use calibration type B to solve for frequency-dependent effects within each spectral window).

To solve for G on, say, fields 1 & 2, on a 90s timescale, and do so relative to gaincurve corrections:

```
gaincal('data.ms',
        caltable='cal.G',          # Write solutions to disk file 'cal.G'
        field='0,1',              # Restrict field selection
        solint=90.0,              # Solve for phase and amp on a 90s timescale
        gaintable=['cal.gc']      # a gain curve table from gencal
        refant='3')               #

plotcal('cal.G','amp')          # Inspect solutions
```

These G solution will be referenced to antenna 4. Choose a well-behaved antenna that is located near the center of the array and is ever-present for the reference antenna. For non-polarization datasets, reference antennas need not be specified although you can if you want. If no reference antenna is specified, an effective phase reference that is an average over the data will be calculated and used. For data that requires polarization calibration, you must choose a reference antenna that has a constant phase difference between the right and left polarizations (e.g. no phase jumps or drifts). If no reference antenna (or a poor one) is specified, the phase reference may have jumps in the R-L phase, and the resulting polarization angle response will vary during the observation, thus corrupting the polarization imaging.

To apply this solution, along with the gain curve correction, to the calibrators (fields 0,1) and the target source (field 2):

```
applycal('data.ms',
         field='0,1,2',          # Restrict field selection (cals + src)
         gaintable=['cal.gc','cal.G']) # Apply gc and G solutions to correct data
```

The calibrated data is written to the CORRECTED_DATA column, with `calwt=True` by default. This parameter can also be a list of Boolean values for which each entry then controls the calculation of weights based on each individual input calibration table. `calwt=False` will recompute the weights from the SIGMA column, thus resetting the weights to their original value.

Alert: *Current* (as of February 2014) Jansky VLA data has no calibrated weights (unless they are computed from switched power calibration). To avoid trouble, `calwt=False` should be set for those data sets. Older, pre-upgrade VLA data should still be calibrated with `calwt=True`.

4.4.3.2 Polarization-independent Gain (T)

At high frequencies, it is often the case that the most rapid time-dependent gain errors are introduced by the troposphere, and are polarization-independent. It is therefore unnecessary to solve for separate time-dependent solutions for both polarizations, as is the case for 'G'. Calibration type 'T' is available to calibrate such tropospheric effects, differing from 'G' only in that a single common solution for both polarizations is determined. In cases where only one polarization is observed, type 'T' is adequate to describe the time-dependent complex multiplicative gain calibration.

In the following example, we assume we have a 'G' solution obtained on a longish timescale (longer than a few minutes, say), and we want a residual 'T' solution to track the polarization-independent variations on a very short timescale:

```
gaincal('data.ms',          # Visibility dataset
        caltable='cal.T',   # Specify output table name
        gaintype='T',      # Solve for T
        field='0,1',       # Restrict data selection to calibrators
        solint=3.0,        # Obtain solutions on a 3s timescale
        gaintable='cal120.G') # Pre-apply prior G solution
```

For dual-polarization observations, it will always be necessary to obtain a 'G' solution to account for differences and drifts between the polarizations (which traverse different electronics), but solutions for rapidly varying polarization-independent effects such as those introduced by the troposphere will be optimized by using 'T'. Note that 'T' can be used in this way for self-calibration purposes, too.

4.4.3.3 GSPLINE solutions

At high radio frequencies, where tropospheric phase fluctuates rapidly, it is often the case that there is insufficient signal-to-noise ratio to obtain robust 'G' or 'T' solutions on timescales short enough to track the variation. In this case it is desirable to solve for a best-fit functional form for each antenna using the 'GSPLINE' solver. This fits a time-series of cubic B-splines to the phase and/or amplitude of the calibrator visibilities.

The `combine` parameter (§ 4.4.1.5) can be used to combine data across spectral windows, scans, and fields. Note that if you want to use `combine='field'`, then all fields used to obtain a 'GSPLINE' amplitude solution must have models with accurate relative flux densities. Use of incorrect relative flux densities will introduce spurious variations in the 'GSPLINE' amplitude solution.

The 'GSPLINE' solver requires a number of unique additional parameters, compared to ordinary 'G' and 'T' solving. The sub-parameters are:

```
gaintype      = 'GSPLINE' # Type of solution (G, T, or GSPLINE)
splintime     = 3600.0    # Spline (smooth) timescale (sec), default=1 hours
npointaver   = 3         # Points to average for phase wrap (okay)
phasese-wrap  = 180      # Wrap phase when greater than this (okay)
```

The duration of each spline segment is controlled by `splintime`. The actual `splintime` will be adjusted such that an integral number of equal-length spline segments will fit within the overall range of data.

Phase splines require that cycle ambiguities be resolved prior to the fit; this operation is controlled by `npointaver` and `phaseswap`. The `npointaver` parameter controls how many contiguous points in the time-series are used to predict the cycle ambiguity of the next point in the time-series, and `phaseswap` sets the threshold phase jump (in degrees) that would indicate a cycle slip. Large values of `npointaver` improve the SNR of the cycle estimate, but tend to frustrate ambiguity detection if the phase rates are large. The `phaseswap` parameter may be adjusted to influence when cycles are detected. Generally speaking, large values ($> 180^\circ$) are useful when SNR is high and phase rates are low. Smaller values for `phaseswap` can force cycle slip detection when low SNR conspires to obscure the jump, but the algorithm becomes significantly less robust. More robust algorithms for phase-tracking are under development (including fringe-fitting).

For example, to solve for 'GSPLINE' phase and amplitudes, with splines of duration 600 seconds,

```
gaincal('data.ms',
        caltable='cal.spline.ap',
        gaintype='GSPLINE'      # Solve for GSPLINE
        calmode='ap'           # Solve for amp & phase
        field='0,1',           # Restrict data selection to calibrators
        splintime=600.)        # Set spline timescale to 10min
```

ALERT: The 'GSPLINE' solutions cannot yet be used in `fluxscale`. You should do at least some 'G' amplitude solutions to establish the flux scale, then do 'GSPLINE' in phase before or after to fix up the short timescale variations. Note that the “phase tracking” algorithm in 'GSPLINE' needs some improvement.

4.4.3.4 Antenna Delays — 'K' solutions

`gaintype='K'` solves for simple antenna-based delays via Fourier transforms of the spectra on baselines to the reference antenna. This is not a global fringe fit but will be useful for deriving delays from data of reasonable snr. If `combine` includes 'spw', multi-band delays solved jointly from all selected spectral windows will be determined, and will be identified with the first spectral window id in the output `caltable`. When applying a multi-band delay table, `spwmap` is required to distribute the solutions to all spectral windows.

After solving for delays, a subsequent `bandpass` is recommended to describe higher-order channel-dependent variation in the phase (and amplitude).

4.4.3.5 Cross-Hand Delays — 'KCROSS' solutions

`gaintype='KCROSS'` solves for a global cross-hand delay. Use `parang=T` and apply prior gain and bandpass solutions. **Alert:** Multi-band delays are not yet supported for `KCROSS` solutions.

4.4.4 Establishing the Flux Density Scale (fluxscale)

The 'G' or 'T' solutions obtained from calibrators for which the flux density was unknown and assumed to be 1 Jansky are correct in a time- and antenna- relative sense, but are mis-scaled by a factor equal to the inverse of the square root of the true flux density. This scaling can be corrected by enforcing the constraint that mean gain amplitudes determined from calibrators of unknown flux density should be the same as determined from those with known flux densities. The `fluxscale` task exists for this purpose.

The inputs for `fluxscale` are:

```
# fluxscale :: Bootstrap the flux density scale from standard calibrators
vis           =      ''      # Name of input visibility file (MS)
caltable     =      ''      # Name of input calibration table
fluxtable    =      ''      # Name of output, flux-scaled calibration table
reference    =      ['']     # Reference field name(s) (transfer flux scale FROM)
transfer     =      ['']     # Transfer field name(s) (transfer flux scale TO), '' ->
                                # all
listfile     =      ''      # Name of listfile that contains the fit information.
                                # Default is (no file).
append      =      False    # Append solutions?
refspwmap    =      [-1]    # Scale across spectral window boundaries. See help
                                # fluxscale
gainthreshold =      -1.0   # Threshold (% deviation from the median) on gain
                                # amplitudes to be used in the flux scale calculation
antenna     =      ''      # antennas to include/exclude
incremental  =      False   # incremental caltable
fitorder    =      1       # order of spectral fitting
display     =      False   # display some statistics of flux scaling
```

Before running `fluxscale`, one must have first run `setjy` for the reference sources and run a `gaincal` that includes `reference` and `transfer` fields. After running `fluxscale` the output `fluxtable` caltable will have been scaled such that the correct scaling will be applied to the `transfer` sources.

For example, given a 'G' table, e.g. 'cal.G', containing solutions for a flux density calibrator (in this case '3C286') and for one or more gain calibrator sources with unknown flux densities (in this example '0234+285' and '0323+022'):

```
fluxscale(vis='data.ms',
          caltable='cal.G',           # Select input table
          fluxtable='cal.Gflx',      # Write scaled solutions to cal.Gflx
          reference='3C286',         # 3C286 = flux calibrator
          transfer='0234+258, 0323+022') # Select calibrators to scale
```

The output table, 'cal.Gflx', contains either the scaling factors alone (`incremental=T`) to be used alongside with the input gain table 'cal.G', or a scaled version of the gain table (`incremental=F`), that replaces it for the execution of `applycal`.

Note that the assertion that the gain solutions are independent of the calibrator includes the assumption that the gain amplitudes are strictly not systematically time-dependent in any way. While synthesis antennas are designed as much as possible to achieve this goal, in practice, a number of effects conspire to frustrate it. When relevant, it is advisable to pre-apply `gaincurve` and `opacity` corrections when solving for the 'G' solutions that will be flux-scaled (see § 4.3 and § 4.4.1.3). When the 'G' solutions are essentially constant for each calibrator separately, the fluxscale operation is likely to be robust.

`fluxscale` will report the fluxes of each spw for each source. In addition, it will attempt a fit across the spws of each source and report a spectral index and curvature ($S \propto (\nu/\nu_0)^{\alpha+\beta*\log(\nu/\nu_0)}$). This information can be subsequently used to build up a model for the spectral slope of a calibrator with the `setjy` task if required.

The `fluxscale` task can be executed on either 'G' or 'T' solutions, but it should only be used on one of these types if solutions exist for both and one was solved relative to the other (use `fluxscale` only on the first of the two).

ALERT: The 'GSPLINE' option is not yet supported in `fluxscale` (see § 4.4.3.3).

If the `reference` and `transfer` fields were observed in different spectral windows, the `refspwmap` parameter may be used to achieve the scaling calculation across spectral window boundaries.

The `refspwmap` parameter functions similarly to the standard `spwmap` parameter (§ 4.4.1.4), and takes a list of indices indicating the spectral window mapping for the reference fields, such that `refspwmap[i]=j` means that reference field amplitudes from spectral window `j` will be used for spectral window `i`.

Note: You should be careful when you have a dataset with spectral windows with different bandwidths, and you have observed the calibrators differently in the different `spw`. The flux-scaling will probably be different in windows with different bandwidths.

For example,

```
fluxscale(vis='data.ms',
          caltable='cal.G',           # Select input table
          fluxtable='cal.Gflx',      # Write scaled solutions to cal.Gflx
          reference='3C286',         # 3C286 = flux calibrator
          transfer='0234+258,0323+022', # Select calibrators to scale
          refspwmap=[0,0,0])        # Use spwid 0 scaling for spwids 1 & 2
```

will use `spw=0` to scale the others, while in

```
fluxscale(vis='data.ms',
          caltable='cal.G',           # Select input table
          fluxtable='cal.Gflx',      # Write scaled solutions to cal.Gflx
          reference='3C286',         # 3C286 = flux calibrator,
          transfer='0234+285, 0323+022', # select calibrators to scale,
          refspwmap=[0,0,1,1])      # select spwids for scaling,
```

the reference amplitudes from spectral window 0 will be used for spectral windows 0 and 1 and reference amplitudes from spectral window 2 will be used for spectral windows 2 and 3.

4.4.4.1 Using Resolved Calibrators

If the flux density calibrator is resolved, the assumption that it is a point source will cause solutions on outlying antennas to be biased in amplitude. In turn, the `fluxscale` step will be biased on these antennas as well. In general, it is best to use model for the calibrator, but if such a model is not available, it is important to limit the solution on the flux density calibrator to only the subset of antennas that have baselines short enough that the point-source assumption is valid. This can be done by using `antenna` and `uvrange` selection when solving for the flux density calibrator. For example, if antennas 1 through 8 are the antennas among which the baselines are short enough that the point-source assumption is valid, and we want to be sure to limit the solutions to the use of baselines shorter than 15000 wavelengths, then we can assemble properly scaled solutions for the other calibrator as follows (note: specifying both an antenna and a `uvrange` constraint prevents inclusion of antennas with only a small number of baselines within the specified `uvrange` from being included in the solution; such antennas will have poorly constrained solutions):

As an example, we first solve for gain solutions for the flux density calibrator (3C286 observed in field 0) using a subset of antennas

```
gaincal(vis='data.ms',
        caltable='cal.G',      # write solutions to cal.G
        field='0'             # Select the flux density calibrator
        selectdata=True,      # Expand other selectors
        antenna='0~7',        # antennas 0-7,
        uvrange='0~15klambda', # limit uvrange to 0-15klambda
        solint=90)            # on 90s timescales, write solutions
                              # to table called cal.G
```

Now solve for other calibrator (0234+285 in field 1) using all antennas (implicitly) and append these solutions to the same table

```
gaincal(vis='data.ms',
        caltable='cal.G',      # write solutions to cal.G
        field='1',
        solint=90,
        append=T)             # Set up to write to the same table
```

Finally, run `fluxscale` to adjust scaling

```
fluxscale(vis='data.ms',
          caltable='cal.G',      # Input table with unscaled cal solutions
          fluxtable='cal.Gflx', # Write scaled solutions to cal.Gflx
          reference='3C286',     # Use 3c286 as ref with limited uvrange
          transfer='0234+285')  # Transfer scaling to 0234+285
```

The `fluxscale` calculation will be performed using only the antennas common to both fields, but the result will be applied to all antennas on the transfer field. Note that one can nominally get by only with the `uvrange` selection, but you may find that you get strange effects from some antennas only having visibilities to a subset of the baselines and thus causing problems in the solving.

4.4.5 Instrumental Polarization Calibration (D,X)

Full support for instrumental polarization calibration for the circular feed basis (e.g., VLA) is provided in CASA. Support for the linear feed basis (e.g., ALMA) is now practical (as of v4.0) and is also described below. The linear feed basis treatment will continue to be expanded and streamlined for the v4.3 release.

The inputs to `polcal` are:

```
# polcal :: Determine instrumental polarization calibrations
vis                =      ''          # Name of input visibility file
caltable           =      ''          # Name of output gain calibration table
field              =      ''          # Select field using field id(s) or field name(s)
spw                =      ''          # Select spectral window/channels
intent             =      ''          # Select observing intent
selectdata         =      True        # Other data selection parameters
    timerange      =      ''          # Select data based on time range
    uvrange        =      ''          # Select data within uvrange (default units meters)
    antenna        =      ''          # Select data based on antenna/baseline
    scan           =      ''          # Scan number range
    observation    =      ''          # Select by observation ID(s)
    msselect       =      ''          # Optional complex data selection (ignore for now)

solint             =      'inf'        # Solution interval
combine            =      'obs,scan'   # Data axes which to combine
    # for solve (obs, scan, spw, and/or field)
preavg             =      300.0        # Pre-averaging interval (sec)
refant             =      ''          # Reference antenna name(s)
minblperant       =      4            # Minimum baselines _per antenna_ required for solve
minsnr             =      3.0         # Reject solutions below this SNR
poltype            =      'D+QU'      # Type of instrumental
    # polarization solution (see help)
smodel             =      []          # Point source Stokes parameters for source model.
append            =      False        # Append solutions to the (existing) table
docallib           =      False        # Use callib or traditional cal apply parameters
    gaintable      =      []          # Gain calibration table(s) to apply
    gainfield      =      []          # Select a subset of calibrators from gaintable(s)
    interp         =      []          # Interpolation mode (in
    # time) to use for each gaintable
    spwmap         =      []          # Spectral windows
    # combinations to form for gaintable(s)
```

The `polcal` task uses many of the standard calibration parameters as described above in § 4.4.1.

The key parameter controlling `polcal` is `poltype`. The choices are:

- 'D' — Solve for instrumental polarization (leakage D-terms), using the transform of an IQU model; requires no parallactic angle coverage, but if the source polarization is non-zero, the gain calibration must have the correct R-L phase registration. (Note: this is unlikely, so just

use 'D+X' to let the position angle registration float.) This will produce a calibration table of type **D**.

'D+X' — Solve for instrumental polarization D-terms and the polarization position angle correction, using the transform of an IQU model; this mode requires at least 2 distinct parallactic angles to separate the net instrumental polarization and the PA. This will produce a calibration table of type 'D'. **ALERT:** no table of type 'X' will be produced, so you must follow this by a run of `polcal` with `polmode='X'` (see below).

'D+QU' — Solve for instrumental polarization and source $Q + iU$; requires at least 3 distinct parallactic angles to separate the net instrumental polarization from the source Q and U. Effectively sets the polarization PA to the value if the R-L phase difference were 0° . This will produce a calibration table of type 'D'.

'X' — Solve only for the position angle correction; best to use this after getting the D-terms from one of the above modes. Requires the observation of a calibrator with known $Q + iU$ (or at least known U/Q). This will produce a calibration table of type 'X'.

'Df11s' — A specialized mode for instrumental polarization solving for the linear feed basis. This will probably be consolidated with other options in a future release.

There are channelized solution modes for the above options. For example, substitute 'Df' for 'D' in the 'D*' modes described above to get a channelized D-term solution; substitute 'Xf' for 'X' to get channelized position angle correction.

ALERT: `polcal` will obtain a separate D-term solution for each `field` supplied to it. This limitation will be relaxed in the future, enabling more sensitive solutions.

4.4.5.1 Heuristics and Strategies for Polarization Calibration

ALERT: This section concentrates on polarization calibration for the circular feed basis. It will be generalized to include the linear feed basis for the v4.3 release. See § 4.4.5.4 for the currently supported processing steps for the linear feed basis.

Fundamentally, with good ordinary gain (and bandpass, if relevant) calibration already in hand, good polarization calibration must deliver both the instrumental polarization and position angle calibration. An unpolarized source can deliver only the first of these, but does not require parallactic angle coverage. A polarized source can only deliver the position angle calibration also if its polarization is known a priori. Sources that are polarized, but with unknown polarization, must always be observed with sufficient parallactic angle coverage, where "sufficient" is determined by SNR and the details of the solving mode.

These principles are stated assuming the instrumental polarization solution is solved using the "linear approximation" where cross-terms in more than a single product of the instrumental or source polarizations are ignored in the Measurement Equation (see § E). A more general non-linearized solution, with sufficient SNR, may enable some relaxation of the requirements indicated here, and modes supporting such an approach are currently under development.

For instrumental polarization calibration, there are 3 types of calibrator choice:

CASA Polarization Calibration Modes

Cal Polarization	Parallactic Angles	model	polmode	Result
unpolarized	any	set $Q = U = 0$	'D' or 'Df'	D-terms only
known non-zero	2+ scans	set Q, U	'D+X' or 'Df+X'	D-terms and PA
unknown	2+ scans	ignored	'D+QU' or 'Df+QU'	D-terms and source

Note that the parallactic angle ranges spanned by the scans in the modes that require this should be large enough to give good separation between the components of the solution. In practice, 60° is a good target.

Each of these solutions should be followed with a 'X' solution on a source with known polarization position angle (and correct $Q + iU$ in the model). **ALERT:** `polmode='D+X'` will soon be enhanced to deliver this automatically.

The `polcal` task will solve for the 'D' or 'X' terms using the model visibilities that are in the model attached to the MS. Calibration of the parallel hands must have already been carried out using `gaincal` and/or `bandpass` in order to align the phases over time and frequency. This calibration must be supplied through the `gaintable` parameters, but any cal-tables to be used in `polcal` must agree (e.g. have been derived from) the data in the `DATA` column and the FT of the model. Thus, for example, one would not use the cal-table produced by `fluxscale` as the rescaled amplitudes would no longer agree with the contents of the model.

Be careful when using resolved calibrators for polarization calibration. A particular problem is if the structure in Q and U is offset from that in I. Use of a point model, or a resolved model for I but point models for Q and U, can lead to errors in the 'X' calibration. Use of a `uvrange` will help here. The use of a full-Stokes model with the correct polarization is the only way to ensure a correct calibration if these offsets are large.

4.4.5.2 A Note on channelized polarization calibration

When your data has more than one channel per spectral window, it is important to note that the calibrator polarization estimate currently assumes the source polarization signal is coherent across each spectral window. In this case, it is important to be sure there is no large cross-hand delay still present in your data. Unless the online system has accounted for cross-hand delays (typically intended, but not always achieved), the gain and bandpass calibration will only correct for parallel-hand delay residuals since the two polarizations are referenced independently. Good gain and bandpass calibration will typically leave a single cross-hand delay (and phase) residual from the reference antenna. Plots of cross-hand phases as a function of frequency for a strongly polarized source (i.e., that dominates the instrumental polarization) will show the cross-hand delay as a phase slope with frequency. This slope will be the same magnitude on all baselines, but with different sign in the two cross-hand correlations. This cross-hand delay can be estimated using the `gaintype='KCROSS'` mode of `gaincal` (in this case, using the strongly polarized source 3C286):

```

default('gaincal')
vis          = 'polcal_20080224.cband.all.ms'
caltable    = 'polcal.xdelcal'
field       = '3C286'
spw         = ''
solint      = 'inf'
combine     = 'scan'
refant      = 'VA15'
smodel      = [1.0,0.11,0.0,0.0]
gaintype    = 'KCROSS'
gaintable   = ['polcal.gcal','polcal.bcal']
gaincal()

```

Note that `smodel` is used to specify that 3C286 is polarized; it is not important to specify this polarization Stokes parameters correctly, as only the delay will be solved for (not any absolute position angle or amplitude scaling). The resulting solution should be carried forward and applied along with the gain (`.gcal`) and bandpass (`.bcal`) solutions in subsequent polarization calibration steps.

4.4.5.3 A Polarization Calibration Example - Circular Feed Basis (e.g., VLA $\nu > 1$ GHz)

In the following example, we do a standard 'D+QU' solution on the bright source BLLac (2202+422) which has been tracked through a range in parallactic angle:

```

default('polcal')
vis          = 'polcal_20080224.cband.all.ms'
caltable    = 'polcal.pcal'
field       = '2202+422'
spw         = ''
solint      = 'inf'
combine     = 'scan'
preavg     = 300.0
refant      = 'VA15'
minsnr     = 3
poltype     = 'D+QU'
gaintable   = ['polcal.gcal','polcal.bcal','polcal.xdelcal']
gainfield   = ['']
polcal()

```

This assumes `setjy` and `gaincal` have already been run. Note that the original gain-calibration table is used in `gaintable` so that what is in the model is in agreement with what is in the `gaintable`, rather than using the table resulting from `fluxscale`.

Now, we need to set the R-L phase using a scan on 3C48 (0137+331):

```

default('polcal')
vis          = 'polcal_20080224.cband.all.ms'

```

```

caltable      = 'polcal.polx'
field        = '0137+331'
refant       = 'VA15'
minsnr       = 3
poltype      = 'X'
smodel       = [1.0,-0.0348,-0.0217,0.0] # the fractional Stokes for 3C48
gaintable    = ['polcal.gcal','polcal.bcal','polcal.xdelcal','polcal.pcal']
polcal()

```

Note that the fractional polarization of 3C48 has been properly specified in `smodel` here.

If, on the other hand, we had a scan on an unpolarized bright source, for example 3C84 (0319+415), we could use this to calibrate the leakages:

```

default('polcal')
vis          = 'polcal_20080224.cband.all.ms'
caltable     = 'polcal.pcal'
field       = '0319+415'
refant      = 'VA15'
poltype     = 'D'
gaintable   = ['polcal.gcal','polcal.bcal','polcal.xdelcal']
polcal()

```

We would then do the 'X' calibration as before (but using this D-table in `gaintable`).

4.4.5.4 A Polarization Calibration Example - Linear Feed Basis (e.g., ALMA, VLA $\nu < 1$ GHz)

CASA v4.0.0 introduces supports for instrumental polarization calibration for the linear feed basis at a level that is now practical for the general user. Some details remain to be implemented with full flexibility, and much of what follows will be streamlined for the v4.1 release.

Calibrating the instrumental polarization for the linear feed basis is somewhat more complicated than the circular feed basis because the polarization effects (source and instrument) appear in all four correlations at first or zeroth order (whereas for circular feeds, the polarization information only enters the parallel hand correlations at second order). As a result, e.g., the time-dependent gain calibration will be distorted by any non-zero source polarization, and some degree of iteration will be required to isolate the gain calibration if the source polarization is not initially known. These complications can actually be used to advantage in solving for the instrumental calibration; in can be shown, for example, that a significantly linearly polarized calibrator enables a better instrumental polarization solution than an unpolarized calibrator.

In the following example, we show the processing steps for calibrating the instrumental polarization using a strongly ($> 5\%$) polarized point-source calibrator (which is also the time-dependent gain calibrator) that has been observed over a range of parallactic angle (a single scan is not sufficient). We assume that we have calibrated the gain, bandpass, and cross-hand delay as described above, and that the gain calibration (`polcal.gcal`) was obtained assuming the calibrator was unpolarized.

First, we import some utility functions from the CASA recipes area:

```
from recipes.almapolhelpers import *
```

Since the gain calibrator was assumed unpolarized, the time-dependent gain solutions contain information about the source polarization. This can be seen by plotting the amp vs. time for this table using `poln='/'`. The antenna-based polarization amplitude ratios will reveal the sinusoidal (in parallactic angle) of the source polarization. Run a utility method (`qufromgain()`) to extract the apparent source polarization estimates for each spw:

```
qu=qufromgain('polcal.gcal')
```

The source polarization reported for all spws should be reasonably consistent. This estimate is not as good as can be obtained from the cross-hands (see below) since it relies on the gain amplitude polarization ratio being stable which may not be precisely true. However, this estimate will be useful in resolving an ambiguity that occurs in the cross-hand estimates.

Next we estimate both the XY-phase offset and source polarization from the cross-hands. The XY-phase offset is a spectral phase-only bandpass relating the X and Y systems of the reference antenna. The cross-hand delay solved for above represents a systematic component (linear phase in frequency). If the XY-phase is solved for in a channel-dependent manner (as below), it is strictly not necessary to have solved for the cross-hand delay above, but it does not hurt (at it allows reasonably coherent channel averages for data examination). The source polarization occurs in the cross-hands as a sinusoidal function of parallactic angle that is common to both cross-hands on all baselines (for a point-source). If the XY-phase bandpass is uniformly zero, then the source linear polarization function will occur entirely in the real part of the cross-hand visibilities. Non-zero XY-phase has the effect of rotating the source linear polarization signature partially into the imaginary part, where circular (and instrumental) polarization occur (cf. the circular feed basis where the cross-hand phase merely rotates the position angle of linear polarization). The following solve averages all baselines together and first solves for a channelized XY-phase (the slope of the source polarization function in the complex plane), then corrects the slope and solves for a channel-averaged source polarization. This calibration is obtained using `gaintype='XYf+QU'` in `gaincal`:

```
default('gaincal')
vis           = 'polcal_linfeed.ms'
caltable     = 'polcal.xy0amb'      # possibly with 180deg ambiguity
field       = '1'                  # the calibrator
solint      = 'inf'
combine     = 'scan'
preavg      = 200.0                # minimal parang change
smodel      = [1,0,1,0]            # non-zero U assumed
gaintype    = 'XYf+QU'
gaintable   = ['polcal.gcal', 'polcal.bcal', 'polcal.xdelcal']
gaincal()
```

Note that we imply non-zero Stokes U in `smodel`; this is to enforce the assumption of non-zero source polarization signature in the cross-hands in the ratio of data and model. This solve will

report the center-channel XY-phase and apparent Q,U for each spw. The Q,U results should be recognizable in comparison to that reported by `qufromgain()` above. However, since the XY-phase has a 180 degree ambiguity (you can rotate the source polarization signature to lie entirely in the visibility real part by rotating clockwise or counter-clockwise), some or all spw QU estimates may have the wrong sign. We correct this using the `xyamb()` utility method, using the `qu` obtained from `qufromgain()` above (which is not ambiguous):

```
S=xyamb(xy='polcal.xy0amb',qu=qu,xyout='polcal.xy0')
```

The python variable `S` now contains the mean source model (Stokes $I = 1$; fractional Q,U; $V=0$) that can be used in a revision of the gain calibration and instrumental polarization calibration.

Next we revise the gain calibration using the full polarization source model:

```
default('gaincal')
vis           = 'polcal_linfeed.ms'
caltable     = 'polcal.gcal1'
field       = '1'
solint      = 'int'           # or whatever was used previously
smodel      = S              # obtained from xyamb
gaintype    = 'G'
gaintable   = ['polcal.bcal']
parang      = T              # so source poln properly rotated
gaincal()
```

Note that `parang=T` so that the supplied source linear polarization is properly rotated in the parallel-hand visibility model. This new gain solution can be plotted with `poln='/'` as above to show that the source polarization is no longer distorting it. Also, if `qufromgain` is run on this new gain table, the reported source polarization should be statistically indistinguishable from zero.

Finally, we can now solve for the instrumental polarization:

```
default('polcal')
vis           = 'polcal_linfeed.ms'
caltable     = 'polcal.dcal'
field       = '1'
solint      = 'inf'
combine     = 'scan'
preavg      = 200
poltype     = 'Dfills'       # freq-dep LLS solver
refant      = ''            # no reference antenna
smodel      = S
gaintable   = ['polcal.gcal1','polcal.bcal','polcal.xdelcal','polcal.xy0']
polcal()
```

Note that no reference antenna is used since this solve will produce an absolute instrumental polarization solution that is registered to the assumed source polarization (`S`) and prior calibrations.

Applying a refant (referring all instrumental polarization terms to a reference antenna's X feed, which would then be assumed perfect) would, in fact, discard valid information about the imperfections in the reference antenna's X feed. (Had we used an unpolarized calibrator, we would not have a valid xy-phase solution, nor would we have had access to the absolute instrumental polarization solution demonstrated here.)

A few points:

- Since the gain, bandpass, and XY-phase calibrations were obtained prior to the instrumental polarization solution and maybe distorted by it, it is generally desirable to resolve for them using the instrumental polarization solution. In effect, this means iterating the sequence of calibration steps using all of the best of the available information at each stage, including the source polarization (and `parang=T`). This is a generalization of traditional self-calibration. For the CASA v4.1 release, we expect to provide utility methods for iteration.
- If the source linear polarization fraction and position angle is known *a priori*, the processing steps outlined above can be amended to use that source polarization assertion in the gain and instrumental calibration solves. The `qufromgain()` method is not needed (but can be used to verify assumptions), the `gaincal(...,gaintype='XYf+QU',...)` should not be altered (parallactic angle coverage is still required!), and the `xyamb()` run should use the *a priori* polarization for `qu`. If there is likely to be a large systematic offset in the mean feed position angle, iteration of the gain, bandpass, and instrumental polarization terms is required to properly isolate the calibration effects.
- Note that the above process does not explicitly include a position angle calibration. In effect, the estimated source polarization sets the mean feed position angle as the reference position angle, and this is usually within a degree or so of optimal. If your mean X feed position angle is not ~ 0 degrees, and your MS does not account for the offset in its FEED subtable, be careful in your interpretation of the final position angle. Currently, the circular feed-specific position angle calibration modes of `polcal (poltype='X' or 'Xf')` will not properly handle the linear feed basis; this will be fixed in the CASA v4.1 release.

A full processing example for linear feed basis polarimetry is under development and will be distributed with an upcoming CASA release.

4.4.6 Baseline-based Calibration (`blcal`)

You can use the `blcal` task to solve for baseline-dependent (non-closing) errors. **WARNING:** this is in general a very dangerous thing to do, since baseline-dependent errors once introduced are difficult to remove. You must be sure you have an excellent model for the source (better than the magnitude of the baseline-dependent errors).

The inputs are (note that `blcal` does not yet use the `docallib` parameter):

```
# blcal :: Calculate a baseline-based calibration solution (gain or bandpass)
vis          =          ''          # Name of input visibility file
```



```

caltable      =      ''      # Name of output gain calibration table
field         =      ''      # Select field using field id(s) or field name(s)
spw           =      ''      # Select spectral window/channels
intent       =      ''      # Select observing intent
selectdata   =      False    # Other data selection parameters
solint       =      'inf'    # Solution interval
combine      =      'scan'   # Data axes which to combine for solve (scan, spw,
# and/or field)
freqdep      =      False    # Solve for frequency dependent solutions
calmode      =      'ap'     # Type of solution" ('ap', 'p', 'a')
solnorm      =      False    # Normalize average solution amplitudes to 1.0
gaintable    =      ['']    # Gain calibration table(s) to apply on the fly
gainfield    =      ['']    # Select a subset of
# calibrators from gaintable(s)
interp       =      ['']    # Interpolation mode (in
# time) to use for each gaintable
spwmap       =      []      # Spectral windows combinations to form for
# gaintable(s)
gaincurve    =      False    # Apply internal VLA antenna
# gain curve correction
opacity      =      []      # Opacity correction to apply (nepers), per spw
parang       =      False    # Apply parallactic angle correction

```

The `freqdep` parameter controls whether `blcal` solves for “gain” (`freqdep=False`) or “bandpass” (`freqdep=True`) style non-closing calibration.

Other parameters are the same as in other calibration tasks. These common calibration parameters are described in § 4.4.1.

4.5 Plotting and Manipulating Calibration Tables

At some point, the user should examine (plotting or listing) the calibration solutions. Calibration tables can also be manipulated in various ways, such as by interpolating between times (and sources), smoothing of solutions, and accumulating various separate calibrations into a single table.

4.5.1 Plotting Calibration Solutions (`plotcal`)

The `plotcal` task is available for examining solutions of all of the basic solvable types (G, T, B, D, M, MF, K). The inputs are:

```

# plotcal :: An all-purpose plotter for calibration results:

caltable     =      ''      # Name of input calibration table
xaxis        =      ''      # Value to plot along x axis (time,chan,amp,phase,real,imag,snr)
yaxis        =      ''      # Value to plot along y axis (amp,phase,real,imag,snr)
poln         =      ''      # Polarization to plot (RL,R,L,XY,X,Y,/)
field        =      ''      # Field names or index: ''=all, '3C286,P1321*', '0~3'

```

```

antenna      =      ''      # Antenna selection.  E.g., antenna='3~5'
spw          =      ''      # Spectral window: ''=all, '0,1' means spw 0 and 1
timerange    =      ''      # Time selection ''=all
subplot      =      111     # Panel number on display screen (yxn)
overplot     =      False   # Overplot solutions on existing display
clearpanel   =      'Auto'  # Specify if old plots are cleared or not
iteration    =      ''      # Iterate on antenna,time,spw,field
plotrange    =      []      # plot axes ranges: [xmin,xmax,ymin,ymax]
showflags    =      False   # If true, show flags
plotsymbol   =      '.'     # pylab plot symbol
plotcolor    =      'blue'  # initial plotting color
markersize  =      5.0     # size of plot symbols
fontsize     =      10.0    # size of label font
showgui      =      True    # Show plot on gui
figfile      =      ''      # ''= no plot hardcopy, otherwise supply name

```

ALERT: Currently, `plotcal` needs to know the MS from which `caltable` was derived to get indexing information. It does this using the name stored inside the table, which does not include the full path, but assumes the MS is in the `cwd`. Thus if you are using a MS in a directory other than the current one, it will not find it. You need to change directories using `cd` in IPython (or `os.chdir()` inside a script) to the MS location.

The controls for the `plotcal` window are the same as for `plotxy` (see § 3.3.2.1).

The `xaxis` and `yaxis` plot options available are:

- `'amp'` — amplitude,
- `'phase'` — phase,
- `'real'` — the real part,
- `'imag'` — the imaginary part,
- `'snr'` — the signal-to-noise ratio,

of the calibration solutions that are in the `caltable`. The `xaxis` choices also include `'time'` and `'channel'` which will be used as the sensible defaults (if `xaxis=''`) for gain and bandpass solutions respectively.

The `poln` parameter determines what polarization or combination of polarization is being plotted. The `poln='RL'` plots both R and L polarizations on the same plot. The respective XY options do equivalent things. The `poln='/'` option plots amplitude ratios or phase differences between whatever polarizations are in the MS (R and L. or X and Y).

The `field`, `spw`, and `antenna` selection parameters are available to obtain plots of subsets of solutions. The syntax for selection is given in § 2.3.

The `subplot` parameter is particularly helpful in making multi-panel plots. The format is `subplot=yxn` where `yxn` is an integer with digit `y` representing the number of plots in the y-axis, digit `x` the number of panels along the x-axis, and digit `n` giving the location of the plot in the panel array (where

$n = 1, \dots, xy$, in order upper left to right, then down). See § 3.3.2.8 for more details on this option.

The `iteration` parameter allows you to select an identifier to iterate over when producing multi-panel plots. The choices for `iteration` are: `'antenna'`, `'time'`, `'spw'`, `'field'`. For example, if per-antenna solution plots are desired, use `iteration='antenna'`. You can then use `subplot` to specify the number of plots to appear on each page. In this case, set the `n` to 1 for `subplot=yxn`. Use the **Next** button on the plotcal window to advance to the next set of plots. Note that if there is more than one timestamp in a 'B' table, the user will be queried to interactively advance the plot to each timestamp, or if `multiplot=True`, the antennas plots will be cycled through for each timestamp in turn. Note that `iteration` can take more than one iteration choice (as a single string containing a comma-separated list of the options). **ALERT:** the iteration order is fixed (independent of the order specified in the `iteration` string), for example:

```
iteration = 'antenna, time, field'
iteration = 'time, antenna, field'
```

will both iterate over each field (fastest) then time (next) and antenna (slowest). The order is:

```
iteration = 'antenna, time, field, spw'
```

from the slowest (outer loop) to fastest (inner loop).

The `markersize` and `fontsize` parameters are especially helpful in making the dot and label sizes appropriate for the plot being made. The screen shots in this section used this feature to make the plots more readable in the cookbook. Adjusting the `fontsize` can be tricky on multi-panel plots, as the labels can run together if too large. You can also help yourself by manually resizing the Plotter window to get better aspect ratios on the plots.

ALERT: Unfortunately, `plotcal` has many of the same problems that `plotxy` does, as they use similar code underneath. An overhaul is underway, so stay tuned.

4.5.1.1 Examples for plotcal

For example, to plot amplitude or phase as a function of time for 'G' solutions (after rescaling by `fluxscale` can look like

```
default('plotcal')
fontsize = 14.0    # Make labels larger
markersize = 10.0 # Make dots bigger

caltable = 'ngc5921.usecase.fluxscale'
yaxis = 'amp'
subplot = 211
plotcal()

yaxis = 'phase'
subplot = 212
plotcal()
```

The results are shown in Figure 4.4. This makes use of the `subplot` option to make multi-panel displays.

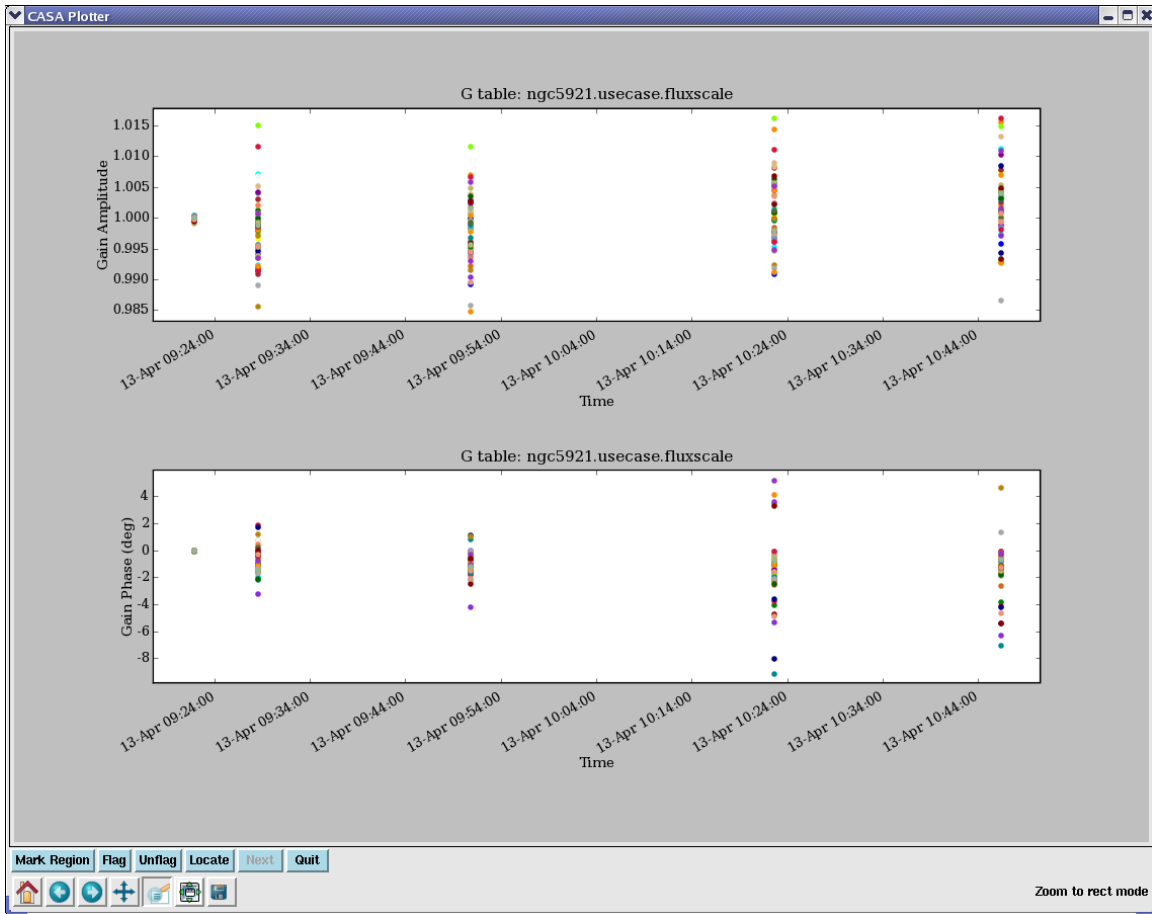


Figure 4.4: Display of the amplitude (upper) and phase (lower) gain solutions for all antennas and polarizations in the `ngc5921` post-fluxscale table.

Similarly, to plot amplitude or phase as a function of channel for 'B' solutions for NGC5921:

```
default('plotcal')
fontsize = 14.0      # Make labels larger
markersize = 10.0  # Make dots bigger

caltable = 'ngc5921.usecase.bcal'
antenna = '1'
yaxis = 'amp'
subplot = 311
plotcal()

yaxis = 'phase'
```

```

subplot = 312
plotcal()

yaxis = 'snr'
subplot = 313
plotcal()

```

The results are shown in Figure 4.5. This stacks three panels with amplitude, phase, and signal-to-noise ratio. We have picked `antenna='1'` to show.

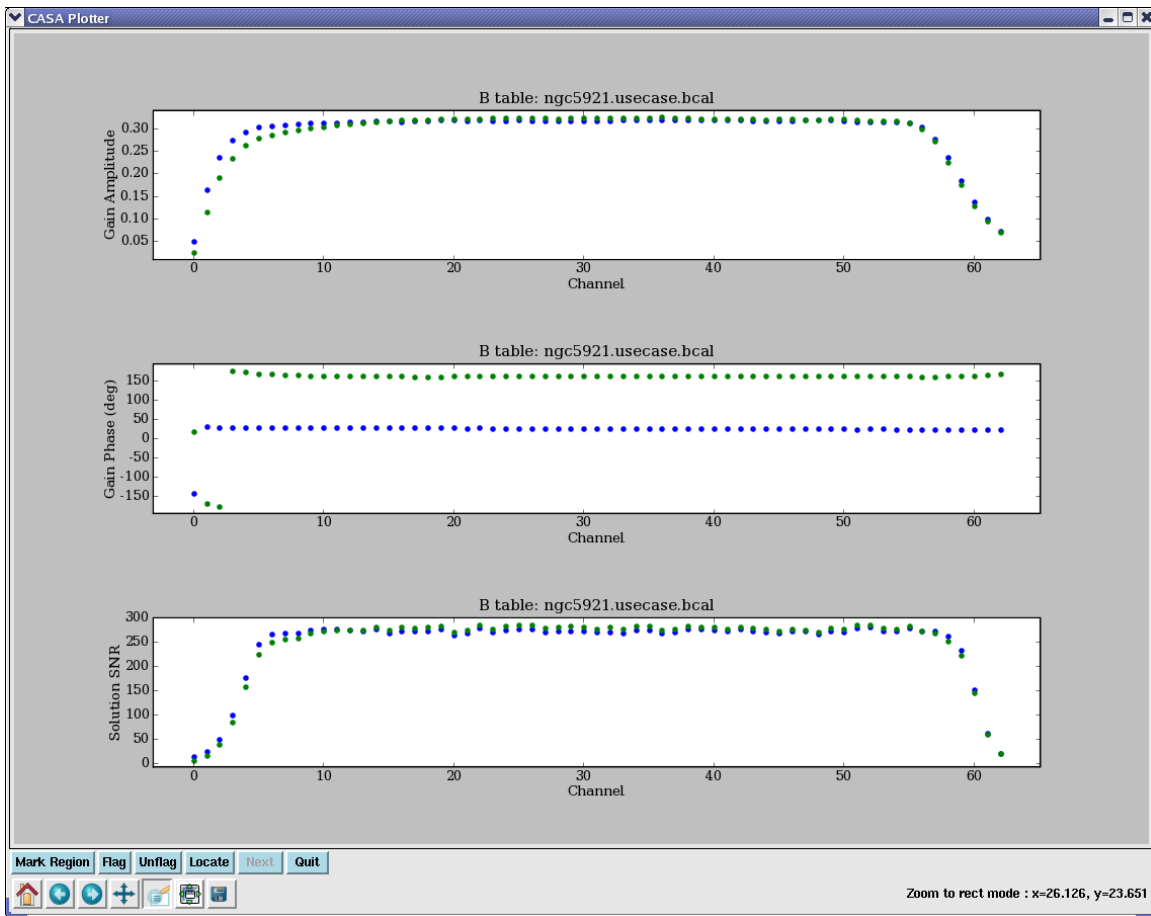


Figure 4.5: Display of the amplitude (upper), phase (middle), and signal-to-noise ratio (lower) of the bandpass 'B' solutions for `antenna='0'` and both polarizations for `ngc5921`. Note the falloff of the SNR at the band edges in the lower panel.

For example, to show 6 plots per page of 'B' amplitudes on a 3×2 grid:

```

default('plotcal')
fontsize = 12.0      # Make labels just large enough

```

```

markersize = 10.0 # Make dots bigger

caltable = 'ngc5921.usecase.bcal'
yaxis = 'amp'
subplot = 231
iteration = 'antenna'

plotcal()

```

See Figure 4.6 for this example. This uses the `iteration` parameter.

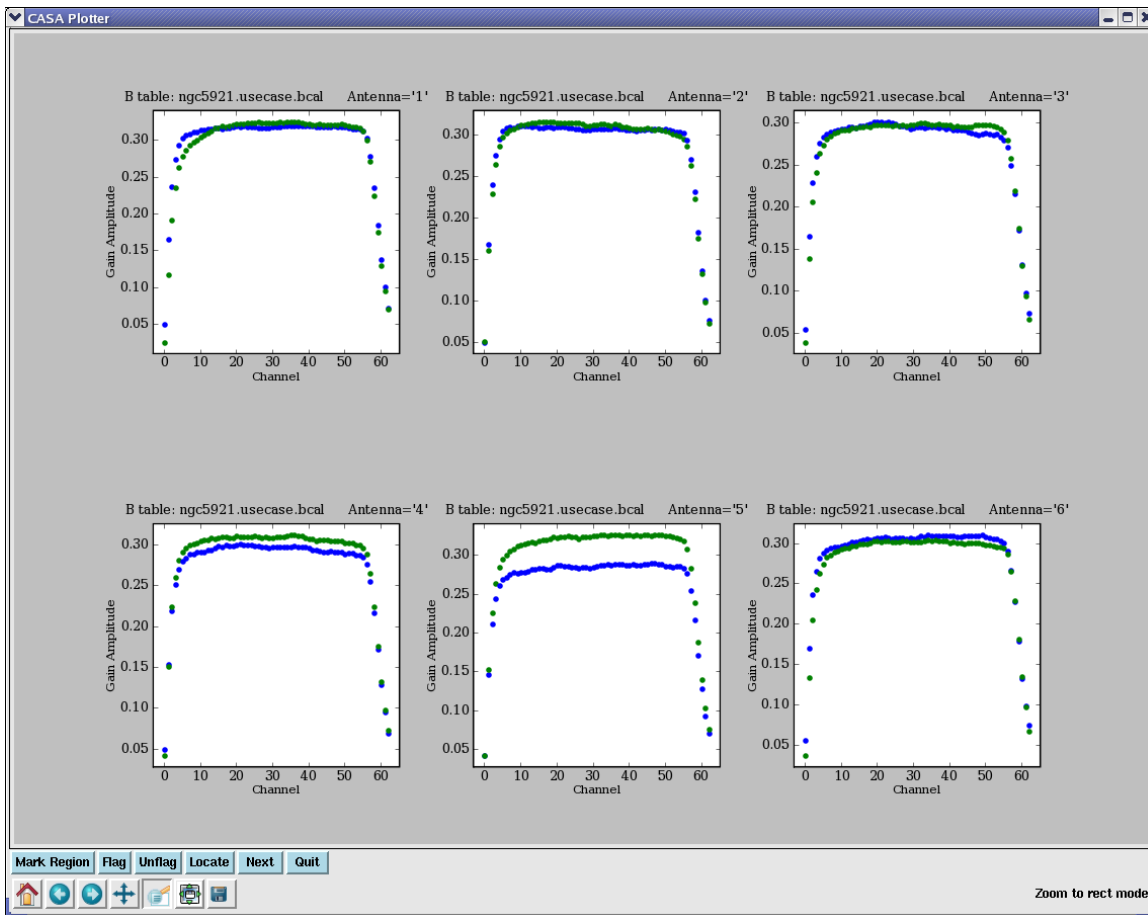


Figure 4.6: Display of the amplitude of the `bandpass` 'B' solutions. Iteration over antennas was turned on using `iteration='antenna'`. The first page is shown. The user would use the `Next` button to advance to the next set of antennas.

4.5.2 Listing calibration solutions with (`listcal`)

The `listcal` task will list the solutions in a specified calibration table.

4.5.3 Calibration table statistics (calstat)

The `calstat` task will print the statistics of solutions in a specified calibration table.

The inputs are:

```
# calstat :: Displays statistical information on a calibration table
caltable      =      ''      # Name of input calibration table
axis          =      'amp'    # Which values to use
      datacolumn =      'gain'  # Which data column to use

useflags      =      True     # Take flagging into account?
      # (not implemented)
```

For example:

```
CASA <3>: calstat('ngc5921.demo.gcal',axis='amp',datacolumn='gain')
```

```
Out [3]:
```

```
{'GAIN': {'max': 1.6031942367553711,
          'mean': 1.4448433067117419,
          'medabsdevmed': 0.0086394548416137695,
          'median': 1.5732669830322266,
          'min': 0.99916577339172363,
          'npts': 280.0,
          'quartile': 0.020265340805053711,
          'rms': 1.4650156497955322,
          'stddev': 0.24271160321065546,
          'sum': 404.55612587928772,
          'sumsq': 600.95579999685287,
          'var': 0.058908922333086665}}
```

```
CASA <4>: calstat('ngc5921.demo.gcal',axis='phase',datacolumn='gain')
```

```
Out [4]:
```

```
{'GAIN': {'max': 0.091214209794998169,
          'mean': -0.015221830284565011,
          'medabsdevmed': 0.012778861448168755,
          'median': -0.012778861448168755,
          'min': -0.15903720259666443,
          'npts': 280.0,
          'quartile': 0.02537553571164608,
          'rms': 0.031241731718182564,
          'stddev': 0.027331476552707856,
          'sum': -4.2621124796782031,
          'sumsq': 0.27329283416317834,
          'var': 0.00074700961055121926}}
```

The statistics can be captured as return variables from the task:

```
CASA <7>: mystat = calstat('ngc5921.demo.gcal',axis='amp',datacolumn='gain')
```



```
CASA <8>: print 'Gain Amp = ',mystat['GAIN']['mean'],'+/-',mystat['GAIN']['stddev']
Gain Amp = 1.44484330671 +/- 0.242711603211
```

ALERT: This task is still under development and currently offers no selection (e.g. by antenna) for the statistics.

4.5.4 Calibration Smoothing (smoothcal)

The `smoothcal` task will smooth calibration solutions (most usefully G or T) over a longer time interval to reduce noise and outliers. The inputs are:

```
# smoothcal :: Smooth calibration solution(s) derived from one or more sources:

vis          =          '' # Name of input visibility file
tablein      =          '' # Input calibration table
caltable     =          '' # Output calibration table
field       =          '' # Field name list
smoothtype   = 'median' # Smoothing filter to use
smoothtime   =          60.0 # Smoothing time (sec)
```

Note that if no `caltable` is specified as output, `smoothcal` will overwrite the input `tablein` calibration table.

The smoothing will use the `smoothtime` and `smoothtype` parameters to determine the new data points which will replace the previous points on the same time sampling grid as for the `tablein` solutions. The currently supported `smoothtype` options:

- `'mean'` — use the mean of the points within the window defined by `smoothtime` (a “boxcar” average),
- `'median'` — use the median of the points within the window defined by `smoothtime` (most useful when many points lie in the interval).

Note that `smoothtime` defines the width of the time window that is used for the smoothing.

ALERT: Note that `smoothcal` currently smooths by `field` and `spw`, and thus you cannot smooth solutions from different sources or bands together into one solution.

An example using the `smoothcal` task to smooth an existing table:

```
smoothcal('n4826_16apr.ms',
          tablein='n4826_16apr.gcal',
          caltable='n4826_16apr.smoothcal',
          smoothtime=7200.,
          smoothtype='mean')

# Plot up before and after tables
plotcal('n4826_16apr.gcal', '', 'amp', antenna='1', subplot=211)
plotcal('n4826_16apr.smoothcal', '', 'amp', antenna='1', subplot=212)
```

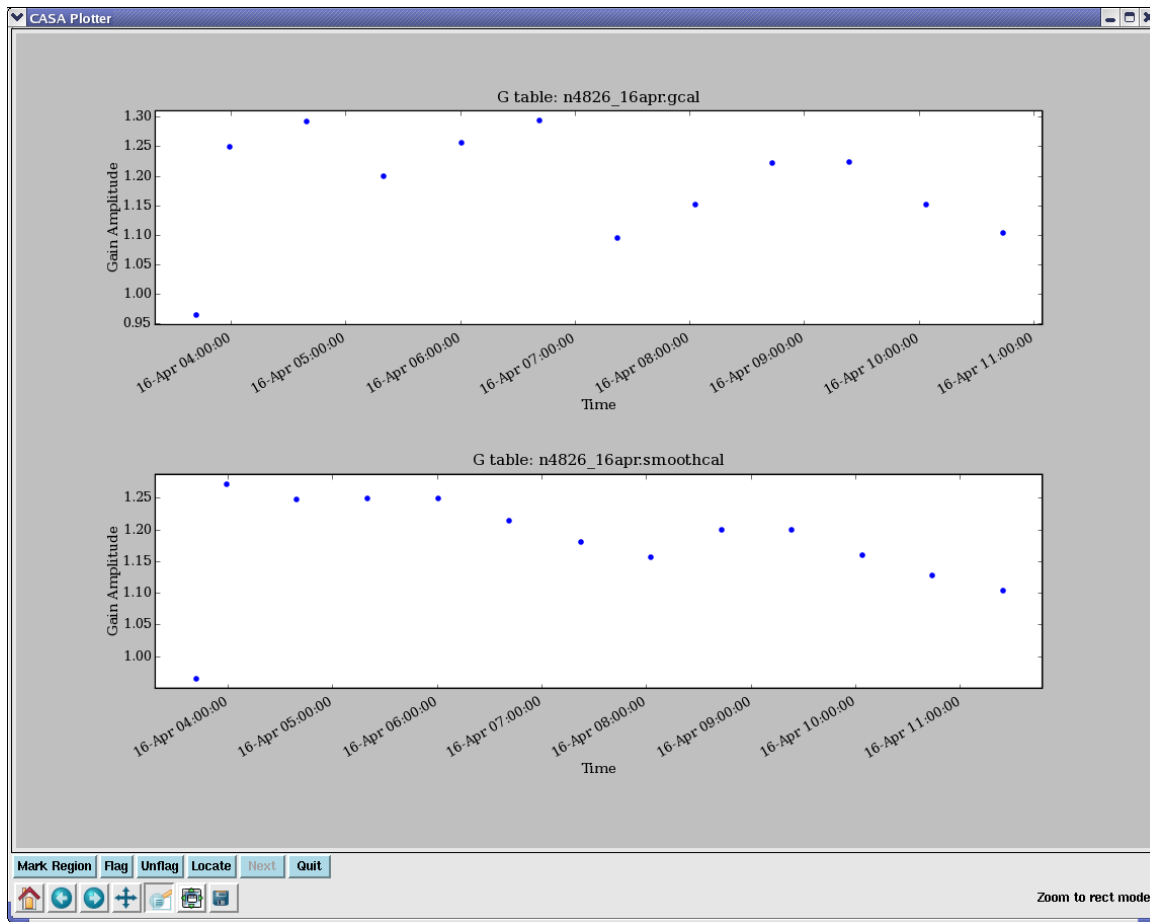


Figure 4.7: The 'amp' of gain solutions for NGC4826 before (top) and after (bottom) smoothing with a 7200 sec `smoothtime` and `smoothtype='mean'`. Note that the first solution is in a different `spw` and on a different source, and is not smoothed together with the subsequent solutions.

This example uses 2 hours (7200 sec) for the smoothing time and `smoothtype='mean'`. The `plotcal` results are shown in Figure 4.7.

4.5.5 Calibration Interpolation and Accumulation (`accum`)

ALERT: The `accum` task is generally no longer recommended for most calibration scenarios. Please write to the NRAO CASA helpdesk if you need support using `accum`.

The `accum` task is used to interpolate calibration solutions onto a different time grid, and to *accumulate* incremental calibrations into a *cumulative* calibration table. The manual accumulation of calibration is rarely required and can usually be achieved implicitly simply by running `applycal` with all the calibration tables given as a list in the `gaintable` parameter (and using `gainfield`,

`spwmap`, and `interp` appropriately. However, sometimes it is desirable to see the interpolated calibration prior to application, and this section describes how this can be done.

Its inputs are:

```
# accum :: Accumulate incremental calibration solutions

vis          =      '' # Name of input visibility file
tablein      =      '' # Input (cumulative) calibration table; use '' on first run
    accumtime =      1.0 # Timescale on which to create cumulative table

incrtable    =      '' # Input incremental calibration table to add
caltable     =      '' # Output (cumulative) calibration table
field        =      '' # List of field names to process from tablein.
calfield     =      '' # List of field names to use from incrtable.
interp       =      'linear' # Interpolation mode to use for resampling incrtable solutions
spwmap       =      [-1] # Spectral window combinations to apply
```

The *mapping* implied here is

```
tablein + incrtable => caltable
```

(mathematically the cal solutions are multiplied as complex numbers as per the Measurement Equation). The `tablein` is optional (see below). You must specify an `incrtable` and a `caltable`.

The `tablein` parameter is used to specify the existing cumulative calibration table to which an incremental table is to be applied. Initially, no such table exists, and if `tablein=''` then `accumulate` will generate one from scratch (on-the-fly), using the timescale (in seconds) specified by the sub-parameter `accumtime`. These nominal solutions will be unit-amplitude, zero-phase calibration, ready to be adjusted by accumulation according to the settings of other parameters. When `accumtime` is negative (the default), the table name specified in `tablein` must exist and will be used. If `tablein` is specified, then the entries in that table will be used.

The `incrtable` parameter is used to specify the incremental table that should be applied to `tablein`. The calibration type of `incrtable` sets the type assumed in the operation, so `tablein` (if specified) must be of the same type. If it is not, `accum` will exit with an error message. (Certain combinations of types and subtypes will be supported by `accum` in the future.)

The `caltable` parameter is used to specify the name of the output table to write. If un-specified (''), then `tablein` will be overwritten. Use this feature with care, since an error here will require building up the cumulative table from the most recent distinct version (if any).

The `field` parameter specifies those field names in `tablein` to which the incremental solution should be applied. The solutions for other fields will be passed to `caltable` unaltered. If the cumulative table was created from scratch in this run of `accumulate`, then the solutions for these other fields will be unit-amplitude, zero-phase, as described above.

The `calfield` parameter is used to specify the fields to select from `incrtable` to use when applying to `tablein`. Together, use of `field` and `calfield` permit completely flexible combinations of

calibration accumulation with respect to fields. Multiple runs of `accum` can be used to generate a single table with many combinations. In future, a `'self'` mode will be enabled that will simplify the accumulation of field-specific solutions.

The `spwmap` parameter gives the mapping of the spectral windows in the `incrtable` onto those in `tablein` and `caltable`. The syntax is described in § 4.4.1.4.

The `interp` parameter controls the method used for interpolation. The options are (currently): `'nearest'` and `'linear'` for time-dependent interpolation, and `'nearest'`, `'linear'`, `cubic`, and `spline` for (optional) frequency-dependent interpolation. These are described in § 4.4.1.4. For most purposes, the `'linear'` option should suffice.

We now describe the two uses of `accum`.

4.5.5.1 Interpolation using (`accum`)

ALERT: The `accum` task is generally no longer recommended for most calibration scenarios. Please write to the NRAO CASA helpdesk if you need support using `accum`.

Calibration solutions (most notably G or T) can be interpolated onto the timestamps of the science target observations using `accum`.

The following example uses `accum` to interpolate an existing table onto a new time grid:

```
accum(vis='n4826_16apr.ms',
      tablein='',
      accumtime=20.0,
      incrtable='n4826_16apr.gcal',
      caltable='n4826_16apr.20s.gcal',
      interp='linear',
      spwmap=[0,1,1,1,1,1])

plotcal('n4826_16apr.gcal', '', 'phase', antenna='1', subplot=211)
plotcal('n4826_16apr.20s.gcal', '', 'phase', antenna='1', subplot=212)
```

See Figure 4.8 for the `plotcal` results. The data used in this example is BIMA data (single polarization YY) where the calibrators were observed in single continuum spectral windows (`spw='0,1'`) and the target NGC4826 was observed in 64-channel line windows (`spw='2,3,4,5'`). Thus, it is necessary to use `spwmap=[0,1,1,1,1,1]` to map the bandpass calibrator in `spw='0'` onto itself, and the phase calibrator in `spw='1'` onto the target source in `spw='2,3,4,5'`.

4.5.5.2 Incremental Calibration using (`accum`)

It is occasionally desirable to solve for and apply calibration incrementally. This is the case when a calibration table of a certain type already exists (from a previous solve), a solution *of the same type* and incremental *relative to the first* is required, and it is not possible or convenient to recover the cumulative solution by a single solve.

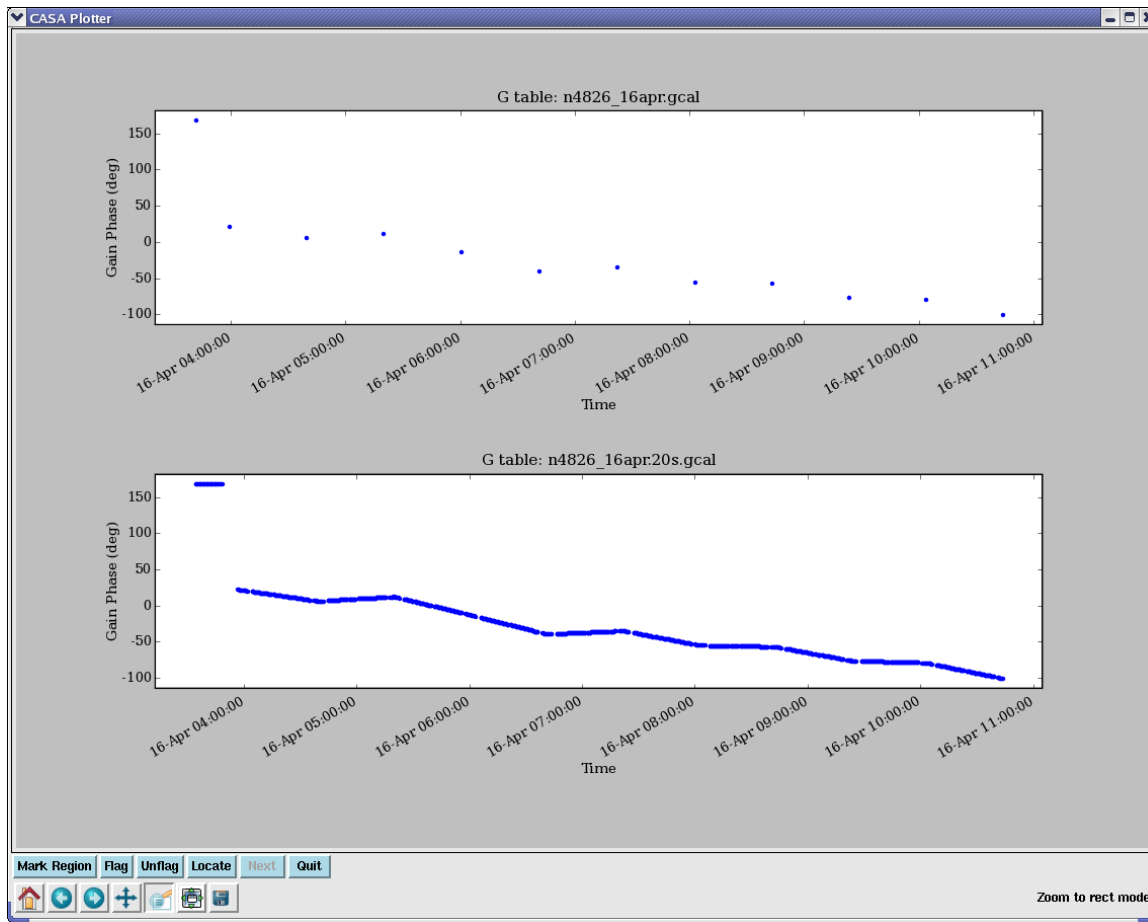


Figure 4.8: The 'phase' of gain solutions for NGC4826 before (top) and after (bottom) 'linear' interpolation onto a 20 sec `accumtime` grid. The first scan was 3C273 in `spw='0'` while the calibrator scans on 1331+305 were in `spw='1'`. The use of `spwmap` was necessary to transfer the interpolation correctly onto the NGC4826 scans.

Much of the time, it is, in fact, possible to recover the cumulative solution. This is because the equation describing the solution for the incremental solution (using the original solution), and that describing the solution for their product are fundamentally the same equation—the cumulative solution, if unique, must always be the same no matter what initial solution is. One circumstance where an incremental solution is necessary is the case of *phase-only* self-calibration relative to a full amplitude and phase calibration already obtained (from a different field).

For example, a phase-only 'G' self-calibration on a target source may be desired to tweak the full amplitude and phase 'G' calibration already obtained from a calibrator. The initial calibration (from the calibrator) contains amplitude information, and so must be carried forward, yet the phase-only solution itself cannot (by definition) recover this information, as a full amplitude and phase self-calibration would. In this case, the initial solution must be applied while solving for the

phase-only solution, then the two solutions combined to form a cumulative calibration embodying the net effect of both. In terms of the Measurement Equation, the net calibration is the product of the initial and incremental solutions.

Cumulative calibration tables also provide a means of generating carefully interpolated calibration, on variable user-defined timescales, that can be examined prior to application to the data with `applycal`. The solutions for different fields and/or spectral windows can be interpolated in different ways, with all solutions stored in the same table.

The only difference between incremental and cumulative calibration tables is that incremental tables are generated directly from the calibration solving tasks (`gaincal`, `bandpass`, etc.), and cumulative tables are generated from other cumulative and incremental tables via `accum`. In all other respects (internal format, application to data with `applycal`, plotting with `plotcal`, etc.), they are the same, and therefore interchangeable. Thus, `accumulate` and cumulative calibration tables need only be used when circumstances require it.

Other Packages:

The analog of `accum` in classic AIPS is the use of `CLCAL` to combine a series of (incremental) SN calibration tables to form successive (cumulative) CL calibration tables. AIPS SN/CL tables are the analog of 'G' tables in CASA.

The `accum` task represents a generalization on the classic AIPS `CLCAL` (see sidebox) model of cumulative calibration in that its application is not limited to accumulation of 'G' solutions. In principle, any basic calibration type can be accumulated (onto itself), as long as the result of the accumulation (matrix product) is of the same type. This is true of all the basic types, except 'D'. Accumulation is currently supported for 'B', 'G', and 'T', and, in future, 'F' (ionospheric Faraday rotation), delay-rate, and perhaps others. Accumulation of certain specialized types (e.g., 'GSPLINE', 'TOPAC', etc.) onto the basic types will be supported in the near future. The treatment of various calibration from ancillary data (e.g., system temperatures, weather data, WVR, etc.), as they become available, will also make use of `accumulate` to achieve the net calibration.

Note that accumulation only makes sense if treatment of a uniquely incremental solution is required (as described above), or if a careful interpolation or sampling of a solution is desired. In all other cases, re-solving for the type in question will suffice to form the net calibration of that type. For example, the product of an existing 'G' solution and an amplitude and phase 'G' self-cal (solved with the existing solution applied), is equivalent to full amplitude and phase 'G' self-cal (with no prior solution applied), as long as the timescale of this solution is at least as short as that of the existing solution.

One obvious application is to calibrate the amplitudes and phases on different timescales during self-calibration. Here is an example:

```
# Add clean model
ft(vis='jupiter6cm.usecase.split.ms',
    model='jupiter6cm.usecase.clean1.model')

# Phase only self-cal on 10s timescales
gaincal(vis='jupiter6cm.usecase.split.ms',
        caltable='jupiter6cm.usecase.phasecal1',
        gaintype='G',
```

```

    calmode='p',
    refant='6',
    solint=10.0,
    minsnr=1.0)

# Plot up solution phase and SNR
plotcal('jupiter6cm.usecase.phasecal1','','phase',antenna='1',subplot=211)
plotcal('jupiter6cm.usecase.phasecal1','','snr',antenna='1',subplot=212)

# Amplitude and phase self-cal on scans
gaincal(vis='jupiter6cm.usecase.split.ms',
        caltable='jupiter6cm.usecase.scancal1',
        gaintable='jupiter6cm.usecase.phasecal1',
        gaintype='G',
        calmode='ap',
        refant='6',
        solint='inf',
        minsnr=1.0)

# Plot up solution amp and SNR
plotcal('jupiter6cm.usecase.scancal1','','amp',antenna='1',subplot=211)
plotcal('jupiter6cm.usecase.scancal1','','snr',antenna='1',subplot=212)

# Now accumulate these - they will be on the 10s grid
accum(vis='jupiter6cm.usecase.split.ms',
      tablein='jupiter6cm.usecase.phasecal1',
      incrtable='jupiter6cm.usecase.scancal1',
      caltable='jupiter6cm.usecase.selfcal1',
      interp='linear')

# Plot this up
plotcal('jupiter6cm.usecase.selfcal1','','amp',antenna='1',subplot=211)
plotcal('jupiter6cm.usecase.selfcal1','','phase',antenna='1',subplot=212)

```

The final plot is shown in Figure 4.9

ALERT: Only interpolation is offered in `accum`, no smoothing (as in `smoothcal`).

4.6 Application of Calibration to the Data

After the calibration solutions are computed and written to one or more calibration tables, one then needs to apply them to the data.

4.6.1 Application of Calibration (`applycal`)

Alert: This section is written using the traditional `applycal` parameters. Users are encouraged to consult Appendix G for information on how to use the “Cal Library” to manage and apply

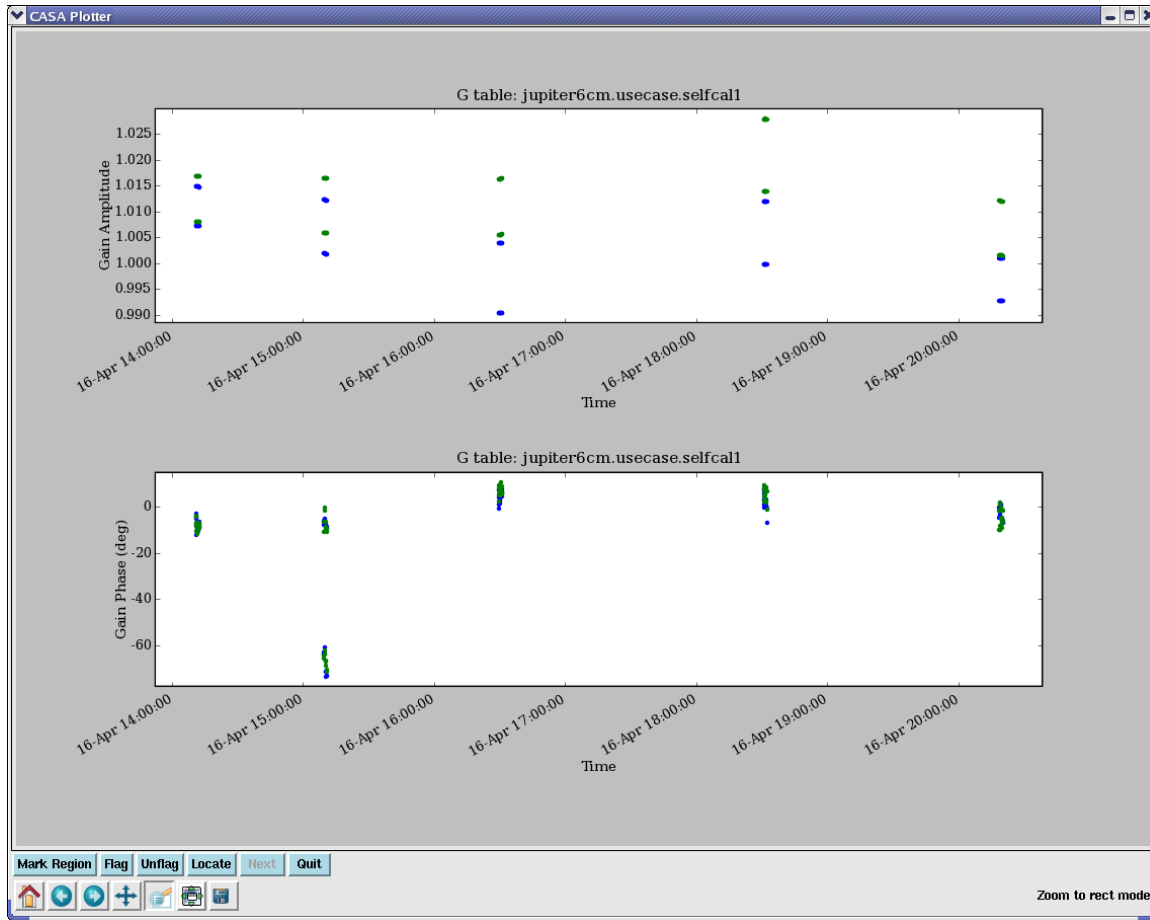


Figure 4.9: The final 'amp' (top) and 'phase' (bottom) of the self-calibration gain solutions for Jupiter. An initial phase calibration on 10s `solint` was followed by an incremental gain solution on each scan. These were accumulated into the cumulative solution shown here.

calibration, which will ultimately provide more flexibility in calibration application, including in a wider variety of applications.

After all relevant calibration types have been determined, they must be applied to the target source(s) before splitting off to a new MS or before imaging. This is currently done by explicitly taking the data in the `DATA` column in the `MAIN` table of the MS, applying the relevant calibration tables, and creating the `CORRECTED_DATA` scratch column. The original `DATA` column is untouched.

The `applycal` task does this. The inputs are:

```
# applycal :: Apply calibrations solutions(s) to data
vis          =      ''      # Name of input visibility file
field        =      ''      # Select field using field id(s) or field name(s)
spw          =      ''      # Select spectral window/channels
```



```

intent          =      ''          # Select observing intent
selectdata      =      True        # Other data selection parameters
  timerange     =      ''          # Select data based on time range
  uvrange       =      ''          # Select data within uvrange (default units meters)
  antenna       =      ''          # Select data based on antenna/baseline
  scan          =      ''          # Scan number range
  observation    =      ''          # Select by observation ID(s)
  msselect      =      ''          # Optional complex data selection (ignore for now)

docallib        =      False       # Use callib or traditional cal apply parameters
  gaintable     =      []          # Gain calibration table(s) to apply on the fly
  gainfield     =      []          # Select a subset of calibrators from gaintable(s)
  interp        =      []          # Interp type in time[,freq],
  # per gaintable. default=linear,linear
  spwmap        =      []          # Spectral windows
  # combinations to form for gaintable(s)
  calwt         =      [True]      # Calibrate data weights per gaintable.

parang          =      False       # Apply parallactic angle correction
applymode       =      ''          # Calibration mode:
  # ""="calflag","trial","flagonly", or "calonly"
flagbackup      =      True        # Automatically back up the
  # state of flags before the run?

```

As in other tasks, setting `selectdata=True` will open up the other selection sub-parameters (see § 2.3). In addition, you can also select data based on the scan intents that were set during the observations (find them through `listobs`). Many of the other parameters are the common calibration parameters that are described in § 4.4.1.

The single non-standard parameter is the `calwt` option to toggle the ability to scale the visibility weights by the inverse of the products of the scale factors applied to the amplitude of the antenna gains (for the pair of antennas of a given visibility). This should in *almost all cases* be set to its default (`True`). The weights should reflect the inverse noise variance of the visibility, and errors in amplitude are usually also in the weights.

Alert: *Current* (as of February 2014) Jansky VLA data has no calibrated weights to the data (unless they are created from switched power). To avoid trouble, `calwt=False` should be set for those data sets. Older, pre-Jansky VLA data should still be calibrated with `calwt=True`.

For `applycal`, the list of final cumulative tables is given in `gaintable`. In this case you will have run `accum` if you have done incremental calibration for any of the types, such as 'G'. You can also feed `gaintable` the full sets and rely on use of `gainfield`, `interp` and `spwmap` to do the correct interpolation and transfer. In particular, for frequency interpolation, the interpolation methods ending in 'PD', `nearestPD` and `linearPD` also scale the phase by the frequency ratio between the measured and interpolated values. It is often more convenient to go through accumulation of each type with `accum` as described above (see § 4.5.5.2), as this makes it easier to keep track of the sequence of incremental calibration as it is solved and applied. You can also do any required smoothing of tables using `smoothcal` (§ 4.5.4), as this is not yet available in `accum` or `applycal`.

`applycal` has different `applymodes`: 'calflag' will apply all flags from a calibration table to the

data and apply the calibration itself to the remaining visibilities. 'trial' will only report on the calibration table flags but not manipulate the data, 'flagonly' applies the flags but not the calibration itself, and 'calonly' will apply the calibration and but not the solution table flags. Data that would 'calflag' would flag are thus passed through uncalibrated. This option can be useful when `applycal` is executed in consecutive steps, one calibration table at a time. Portions of the data that were not calibrated in the first run can then be calibrated in a second run with a different calibration table. This option should be used with care such that no uncalibrated data remains in the final data product.

`applycal` will flag all data that have no calibration solution. Flags will distribute into all of your scratch columns, i.e. it will affect your uncalibrated visibilities, too. To be able to restore the flags to the state before `applycal` is starting its duty, the task will make a backup of your current flags by default (`flagbackup=True`). Restore them with `flagmanager`, if you are not happy with the `applycal` results.

If you are not doing polarization calibration or imaging, then you can set `parang=False` to make the calculations faster. If you are applying polarization calibration, or wish to make polarization images, then set `parang=True` so that the parallactic angle rotation is applied to the appropriate correlations. Currently, you must do this in `applycal` as this cannot be done on-the-fly in `clean` or `mosaic`. See § 4.4.1.3 for more on `parang`.

For example, to apply the final bandpass and flux-scaled gain calibration tables solutions to the NGC5921 data:

```
default('applycal')

vis='ngc5921.usecase.ms'

# We want to correct the calibrators using themselves
# and transfer from 1445+099 to itself and the target N5921

# Start with the fluxscale/gain and bandpass tables
gaintable=['ngc5921.usecase.fluxscale','ngc5921.usecase.bcal']

# pick the 1445+099 (field 1) out of the gain table for transfer
# use all of the bandpass table
gainfield = ['1','*']

# interpolation using linear for gain, nearest for bandpass
interp = ['linear','nearest']

# only one spw, do not need mapping
spwmap = []

# all channels, no other selection
spw = ''
selectdata = False

# no prior calibration
gaincurve = False
```

```

opacity = 0.0

# select the fields for 1445+099 and N5921 (fields 1 and 2)
field = '1,2'

applycal()

# Now for completeness apply 1331+305 (field 0) to itself

field = '0'
gainfield = ['0','*']

applycal()

# The CORRECTED_DATA column now contains the calibrated visibilities

```

In another example, we apply the final cumulative self-calibration of the Jupiter continuum data obtained in the example of § 4.5.5.2:

```

applycal(vis='jupiter6cm.usecase.split.ms',
         gaintable='jupiter6cm.usecase.selfcal1',
         selectdata=False)

```

Again, it is important to remember the relative nature of each calibration term. A term solved for in the presence of others is, in effect, residual to the others, and so must be used in combination with them (or new versions of them) in subsequent processing. At the same time, it is important to avoid isolating the same calibration effects in more than one term, e.g., by solving for both 'G' and 'T' separately (without applying the other), and then using them together.

It is always a good idea to examine the corrected data after calibration (using `plotxy` to compare the raw ('data') and corrected ('corrected') visibilities), as we describe next.

4.6.2 Examine the Calibrated Data

Once the source data is calibrated using `applycal`, you should examine the *uv* data and flag anything that looks bad. If you find source data that has not been flanked by calibration scans, delete it (it will not be calibrated).

For example, to look at the calibrated Jupiter data in the last example given in the previous section:

```

plotxy('jupiter6cm.usecase.split.ms', 'uvdist', 'amp', 'corrected',
       selectdata=True, correlation='RR LL', fontsize = 14.0)

```

will show the `CORRECTED_DATA` column. See Figure 4.10.

See § 3.3 for a description of how to display and edit data using `plotms` or `plotxy`, and § 7.5 for use of the viewer to visualize and edit a Measurement Set.

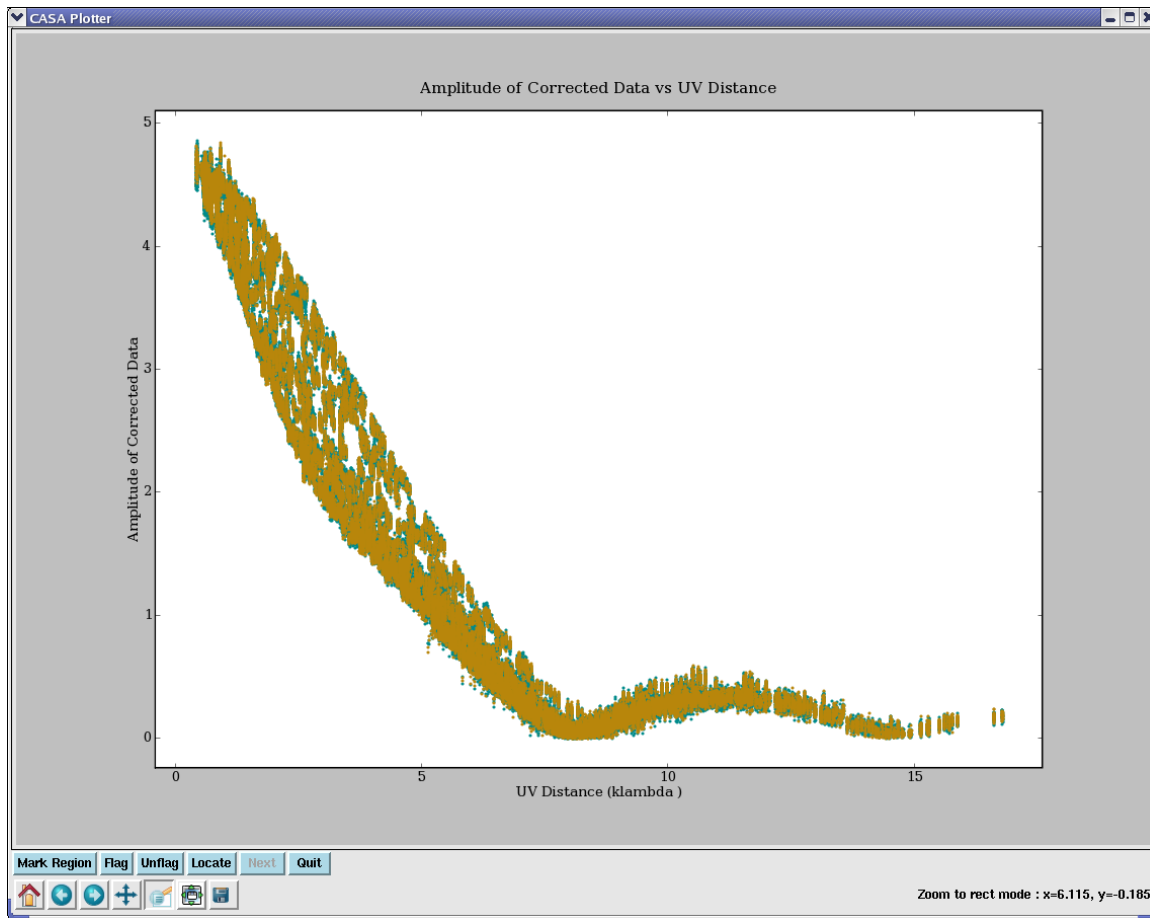


Figure 4.10: The final 'amp' versus 'uvdist' plot of the self-calibrated Jupiter data, as shown in plotxy. The 'RR LL' correlations are selected. No outliers that need flagging are seen.

4.6.3 Resetting the Calibration Models (delmod and clearcal)

Whenever calibration tasks are run, the models associated with the MS will be overwritten. Sometimes, however, one would like to completely remove the model and the task `delmod` can perform this functionality:

```
# delmod :: Deletes model representations in the MS
vis          =      ''          # Name of input visibility file (MS)
otf          =      True        # Delete the on-the-fly model data keywords
  field      =      ''          # Select field using field id(s) or field name(s)

scr          =      False       # Delete the MODEL_DATA scr col (if it exists)
```

To do so, the parameter `otf` should be set to `True`. `delmod` can also be used if for any reason a `MODEL` column was created and should be removed to avoid confusion between the on-the-fly model and the `MODEL` column

(the MODEL_DATA column was required in CASA 3.3 and earlier). This can be achieved with `scr=T`.

`delmod` generally replaces the functionality of the older `clearcal` task. If one still decides to use the MODEL_DATA columns, however, `clearcal` is still useful and will reset both the MODEL_DATA and CORRECTED_DATA columns to DATA:

```
CASA <11>: inp clearcal
# clearcal :: Re-initializes the calibration for a visibility data set
vis          =          ''      # Name of input visibility file (MS)
field        =          ''      # Select field using field id(s) or field name(s)
spw          =          ''      # Select spectral window/channel.
intent       =          ''      # Select observing intent
addmodel     =          False    # Add MODEL_DATA scratch column
```

with `field`, `spw`, and `intent` being data selection parameters. `addmodel` can be used to opt in/out of formation of the MODEL_DATA column.

With the introduction of the on-the-fly calculation of the MODEL visibilities, and the fact that `applycal` overwrites any previously existing CORRECTED_DATA column, `clearcal` is not required anymore unless `usescratch=True` is chosen in calibration tasks, and it is also not recommended to use `clearcal` to create the scratch columns at the beginning of data calibration; all benefits from the on-the-fly model would be made obsolete.

4.7 Other Calibration and UV-Plane Analysis Options

4.7.1 Splitting out Calibrated uv data (split)

The `split` task will apply calibration and output a new sub-MS containing a specified list of sources (usually a single source). The inputs are:

```
# split :: Create a visibility subset from an existing visibility set:
vis          =          ''      # Name of input measurement set
outputvis    =          ''      # Name of output measurement set
datacolumn   = 'corrected'     # Which data column(s) to split out
field        =          ''      # Select field using field id(s) or field name(s)
spw          =          ''      # Select spectral window/channels
width        =          1      # Number of channels to average to form one output channel
antenna      =          ''      # Select data based on antenna/baseline
timebin      =          '0s'    # Value for timeaveraging
timerange    =          ''      # Select data based on time range
scan         =          ''      # Select data based on scan numbers
array        =          ''      # Select (sub)array by array ID number(s)
uvrange      =          ''      # Select data based on uv distance range
```

Usually you will run `split` with `datacolumn='corrected'` as previous operations (e.g. `applycal`) will have placed the calibrated data in the CORRECTED_DATA column of the MS. This will produce a new MS with this corrected data in its DATA column. The modes available in `datacolumn` are:

```
'data', 'model', 'corrected',           # produce MS with single DATA column
'data,model', 'data,corrected', 'model,corrected', # pairs of columns
'all'                                     # all columns 'data,model,corrected'
```

We recommend sticking to the simple single-column modes (e.g. 'data' or 'corrected') or 'all' if all columns are in the MS. Further processing may get confused by mismatched pairs of columns.

For example, to split out 46 channels (5-50) from spw 1 of our NGC5921 calibrated dataset:

```
split(vis='ngc5921.usecase.ms',
      outputvis='ngc5921.split.ms',
      field='2',           # Output NGC5921 data (field 2)
      spw='0:5~50',       # Select 46 chans from spw 0
      datacolumn='corrected') # Take the calibrated data column
```

4.7.1.1 Averaging in split

Time and channel averaging are available using the `timebin` and `width` parameters.

The `timebin` parameter gives the averaging interval. It takes a quantity, e.g.

```
timebin = '30s'
```

and will combine scans during averaging.

When time averaging, the `ignorables` subparameter can be used to specify that the bins should *not* be split by changes in `SCAN_NUMBER`, `(sub)ARRAY_ID`, and/or `STATE_ID`.

The `width` parameter defines the number of channels to average to form a given output channel. This can be specified globally for all `spw`, e.g.

```
width = 5
```

or specified per `spw`, e.g.

```
width = [2,3]
```

to average 2 channels of 1st spectral window selected and 3 in the second one.

ALERT: When averaging channels split will produce negative channel widths (as reported by `listobs`) if frequency goes down with increasing channel number, whether or not the input channel widths are negative. The *bandwidths* and channel resolutions will still be positive.”

4.7.2 Recalculation of uvw values (fixvis)

Sometimes the u,v,w coordinates of a Measurement Set are not recorded correctly by the correlator. In those cases, it may be necessary to recalculate them based on the antenna positions. `fixvis` will perform this task.

```
# fixvis :: Recalculates (u, v, w) and/or changes Phase Center
vis          =      ''      # Name of the input visibility set.
outputvis    =      ''      # Name of the output visibility set. (Can be the same
                             # as vis.)
field        =      ''      # Fields to operate on. = all.
refcode      =      ''      # reference frame to convert UVW coordinates to
reuse        =      True    # base UVW calculation on the old values?
phasecenter  =      ''      # use this direction as phase center
```

A useful feature of `fixvis` is that it can also change the phase center of a Measurement Set. This can be done with absolute coordinates or using offsets. An example is:

```
fixvis(vis='Moon.ms',outputvis='Moon-fixed.ms',field='Moon', phasedir='J2000 9h25m00s 05d12m00s')
```

that will recalculate the u,v,w coordinates relative to the new phase center for the field 'Moon'.

4.7.3 Hanning smoothing of uv data (hanningsmooth)

For strong spectral line sources (like RFI sources), the Gibbs phenomenon may cause ringing across the frequency channels of an observation. This is called the Gibbs phenomenon and a proven remedy is the Hanning smoothing algorithm. Hanning smoothing is a running mean across the spectral axis with a triangle as a smoothing kernel. The central channel is weighted by 0.5 and the two adjacent channels by 0.25 to preserve the flux. Hanning smoothing significantly reduces Gibbs ringing but there's no gain without a penalty and here it is the loss of a factor of two in spectral resolution.

In CASA, the `hanningsmooth` task will apply Hanning smoothing to a spectral line uv data Measurement Set. The inputs are:

```
# hanningsmooth :: Hanning smooth frequency channel data to remove Gibbs ringing
vis          =      ''      # Name of input visibility file (MS)
datacolumn   =      'all'   # the name of the MS column into which
                             # to write the smoothed data
outputvis    =      ''      # name of the output visibility file
                             # (MS)
```

`hanningsmooth` will operate on the input Measurement Set if no `outputvis` file name is provided. This option will keep the disk usage of large datasets under control. But one should be aware that the data is overwritten. If `outputvis` is provided, the task will copy the input MS to a new file with that name and operate there. The `datacolumn` parameter determines which of the data

columns is to be Hanning smoothed: 'all', 'corrected' or 'data'. 'all' refers to both, the `CORRECTED_DATA` and the `DATA` column. If 'corrected' is specified but does not exist in the MS, `hanningsmooth` will create this column for your convenience.

ALERT: We intend to make the kernel size a user supplied parameter. In the longer term we intend to offer other varieties of spectral smoothing as well.

4.7.4 MSttransform (`msttransform`)

`msttransform` is a multipurpose task that provides all the functionality of `split`, `partition`, `cvel`, `hanningsmooth` and `applycal` with the possibility of applying each of these transformations separately or together in an in-memory pipeline, thus avoiding unnecessary I/O steps. The list of transformations that `msttransform` can apply is as follows:

1. Data selection and re-indexing
2. Data partition (create output Multi-MS)
3. On-the-fly calibration (via “Cal Library”)
4. Time average (weighted and baseline dependent)
5. Channel average (weighted)
6. Hanning smooth
7. Combination of spectral windows
8. Spectral regridding and reference frame transformation
9. Separation of spectral windows

Notice that the order in the above list is not arbitrary. When various transformations are applied on the data using `msttransform` the order in which the transformations are pipe one after the other is the one shown in the above list.

Besides `msttransform` in itself, there are a series of tasks that mimic the old interfaces and are based on the `msttransform` framework: `split2`, `cvel2` and `hanningsmooth2`. These tasks will soon replace the old tasks, and new features like spectral weight handling (see appendix F) are only incorporated in the new tasks.

A complete functionality description is available at the following web site:

<http://www.eso.org/~scastro/ALMA/casa/MST/MSTtransformDocs/MSTtransformDocs.html>

4.7.4.1 Data selection and re-indexing

`mstransform / split2` as the old `split` are able to create a new MS with a specific data selection, for instance splitting a science target. The new MS contains only the selected data and also the subtables are re-generated to contain only the metadata matching the data selection. The details about pure split operation are described in section (§ 4.7.1)

```

vis           =      ''      # Name of input Measurement set or Multi-MS.
outputvis    =      ''      # Name of output Measurement Set or Multi-MS.
tileshape    =      [0]     # List with 1 or 3 elements giving the
                           # tile shape of the disk data columns.
field        =      ''      # Select field using ID(s) or name(s).
spw          =      ''      # Select spectral window/channels.
scan         =      ''      # Select data by scan numbers.
antenna      =      ''      # Select data based on antenna/baseline.
correlation  =      ''      # Correlation: '' ==> all, correlation='XX,YY'.
timerange    =      ''      # Select data by time range.
intent       =      ''      # Select data by scan intent.
array        =      ''      # Select (sub)array(s) by array ID number.
uvrange      =      ''      # Select data by baseline length.
observation  =      ''      # Select by observation ID(s).
feed         =      ''      # Multi-feed numbers: Not yet implemented.
datacolumn   = 'corrected'  # Which data column(s) to process.
keepflags    =      True    # Keep *completely flagged rows* or
                           # drop them from the output.
usewtspectrum =      False  # Create a WEIGHT_SPECTRUM column in the output MS.

```

The new features related with data election and re-indexing contained in `mstransform / split2` but not in the old `split` are the following:

- Spectral weight initialization: `mstransform` can initialize the output `WEIGHT/SIGMA_SPECTRUM` columns by specifying `usewtspectrum = True`. The details about spectral weights initialization are described in section (§ 4.3.1)
- Tile shape specification for the data columns: `mstransform` also allows to specify a custom default tile shape for the output data columns, namely a list of 3 elements specifying the number of correlations, channels and rows to be contained in each tile, for instance `tileshape = [4,128,351]` would specify a tile with shape (4 correlations)x(128 channels)x(351 rows). This can be used to optimize the access pattern of subsequent tasks, for instance imaging tasks.
- Support for SPWs with different sets of correlation products: `mstransform / split2` are both able to work when a given SPW is associated with several correlation products (like in some EVLA correlation setups). This is transparent for the user and simply works by using the `spw` data selection parameter normally. It also works in conjunction with the `polarization` parameter, so for instance if a given MS has separated RR and LL data associated with `spw 0` the following data selection would work flawlessly: `spw = '0' polarization = 'LL'`

- Support for multiple channel selection: Both `mstransform` / `split2` are also capable of working with multiple channel selection. This support also goes transparently for the user, by simply following the SPW syntax specified by the `MSSelection` section (§ 2.3.3.1). For example `spw = '4 7:4 59,8:4 13;18 594'`

4.7.4.2 Data partition

`MSTransform` is the framework used by the `partition` task, and even though it can be used directly to produce an MMS by specifying `createmms = True` it is highly recommended to use directly the `partition` task as explained in chapter (§ 10). Nevertheless, for the sake of completeness this is the list of expandable parameters shown when `createmms = True`

```
createmms          =          True          # Create a multi-MS output
                                     from an input MS.
    separationaxis =          'auto'        # Axis to do parallelization across
                                     (scan,spw,auto)
    numsubms       =              9         # The number of Sub-MSs to create
                                     (auto or any number)
```

- `separationaxis`: Specifies the partition axis, across scan, spw or auto. The default mode is auto, which parts the data across both scan and spw distributing the data of each scan and spw across all sub-MS in the most balanced way possible.
- `numsubms`: When this parameter is assigned to an integer it refers to the number of sub-MSs that the output MMS would have. The default is auto which means producing as many sub-MSs as processing engine/servers (in cluster mode).

4.7.4.3 On-the-fly calibration

As of CASA 4.5 `mstransform` incorporates the possibility of applying on-the-fly (OTF) calibration by specifying `docallib = True`, which in turns allows to specify the “Cal Library” filename (`callib` parameter) whose format is described in Appendix G. This transformation is the first one applied to the data, producing effectively a corrected data column on-the-fly, which can be further transformed.

```
docallib          =          True          # Enable OTF calibration
    callib         =              ''        # Cal Library filename
```

- `callib`: Filename pointing to the calibration specification file whose format is described in Appendix G, where conventions and current limitations are also described.

- The old `split` performs a flat average taking into account only the `FLAG` column, whereas `mstransform / split2` use both `FLAG` and spectral weights (when present), resulting in a weighted average. To be specific `WEIGHT_SPECTRUM` is used when averaging `CORRECTED_DATA`, and `SIGMA_SPECTRUM` is used when averaging the `DATA` column.
- Also `mstransform / split2` are able to transform the input `WEIGHT/SIGMA_SPECTRUM` according to the rules of error propagation that apply to a weighted average, which result in an output weight equals to the sum of the input weights. For a detailed reference see, *Data Reduction and Error Analysis* by Bevington & Robinson (3rd Ed., McGraw Hill, 2003).
- Both `mstransform / split2` drop the last output channel bin when there are not enough contributors to fully populate it. For instance, if the input SPW has 128 channels and `chanbin` is 10, the resulting averaged SPW would have 12 channels and not 13 channels.

4.7.4.6 Hanning smooth

Both `mstransform / hanningsmooth2` are able to perform hannining smooth (in `mstransform` by specifying `hanning = True` and in `hanningsmooth2` by default). The only difference w.r.t. the old `hanningsmooth` is that a new MS is created.

4.7.4.7 Combination of spectral windows

Both `mstransform / cvel2` are able combine SPWs (in `mstransform` by specifying `combinespws = True` and in `cvel2` by default). The algorithm is in general the same as the old `cvel`, however, there are two significant differences in the new framework:

- `mstransform` Is able to only combine SPWs, only regrid each SPW separately or to combine all SPWs and regrid them together. The details about independent regridding operation are explained in the following sections.
- `mstransform / cvel2` Automatically detect combination of SPWs with different exposure, and use the `WEIGHT` column (or `WEIGHT_SPECTRUM` if available) in addition of the geometrical factors to calculate the combined visibility in the overlapping regions.

4.7.4.8 Spectral regridding and reference frame transformation

Both `mstransform / cvel2` are able to perform spectral regridding / reference frame transformation (in `mstransform` by specifying `regridms = True` and in `cvel2` by default). However `mstransform` is able to perform spectral regridding / reference frame transformation on each selected SPW separately, that is w/o combining the selected SPWs. As of CASA 4.5 both algorithms are fully aligned including the latest developments to take into account ephemerides and radial velocity correction.

```

regridms          =          True          # Regrid the MS to a new spw,
                                         # channel structure or frame.
    mode          =    'channel'          # Regridding mode
                                         # (channel/velocity/frequency/channel_b).
    nchan         =          -1           # Number of channels in the output spw (-1=all).
    start         =          0           # First channel to use in the output spw
                                         # (mode-dependant)
    width         =          1           # Number of input channels that are
                                         # used to create an output channel.
    nspw          =          1           # Number of output spws to create in output MS.
    interpolation  =    'linear'          # Spectral interpolation method.
    phasecenter   =          ''          # Image phase center: position or field index.
    restfreq      =          ''          # Rest frequency to use for output.
    outframe      =          ''          # Output reference frame (''=keep input frame).
    veltype       =    'radio'          # Velocity definition.

```

4.7.4.9 Separation of spectral windows

A completely new feature in `mstransform` is the ability to separate an input SPW into several ones, or to combine various input SPWs into a single one with a uniform grid (resolving overlaps/gaps) to then separate it in several output SPWs. This option is activated under the regridding section (therefore by specifying `regridms = True`), together with the `nspw` parameter which when bigger than 1 implies that the input SPW / combination of input SPWs should be separated:

```

regridms          =          True          # Regrid the MS to a new spw,
    nspw          =          1           # Number of output spws to create in output MS.

```

4.7.5 Model subtraction from uv data (uvsub)

The `uvsub` task will subtract the Fourier transform of the associated model of the MS (added to the MS with the tasks `ft` or `setjy`) from that in the `CORRECTED_DATA` column in the input MS and store the result in that same `CORRECTED_DATA` column.

The reverse operation is achieved by specifying `reverse = True`: in that case `uvsub` will add the value of the Fourier transform of the associated model to that in the `CORRECTED_DATA` column in the input MS and store the result in that same `CORRECTED_DATA` column.

The inputs are:

```

# uvsub :: Subtract/add model from/to the corrected visibility data.

vis          =          ''          # Name of input visibility file (MS)
reverse      =          False       # reverse the operation (add rather than subtract)

```

For example:

```
uvsub('ngc5921.split.ms')
```

ALERT: Currently, `uvsub` operates on the `CORRECTED_DATA` column in the MS `vis`. Eventually we will provide the option to write out a new MS.

4.7.6 UV-Plane Continuum Subtraction (`uvcontsub`)

At this point, consider whether you are likely to need continuum subtraction. If there is significant continuum emission present in what is intended as a spectral line observation, continuum subtraction may be desirable. You can estimate and subtract continuum emission in the uv -plane prior to imaging or wait and subtract an estimate of it in the image-plane. Note that neither method is ideal, and the choice depends primarily upon the distribution and strength of the continuum emission. Subtraction in the uv -plane is desirable if continuum emission dominates the source, since deconvolution of the line emission will be more robust if it not subject to the deconvolution errors of the brighter continuum. There is also a performance benefit since the continuum is nearly the same in each channel of the observation, and it is desirable to avoid repeating its deconvolution in each channel. However, doing the continuum estimation in the uv -plane has the serious drawback that interpolating visibilities between channels is only a good approximation for emission from near the phase center. Thus, `uvcontsub` will do an increasingly poor job for emission distributed further from the phase center. If the continuum emission is relatively weak, it is usually adequate to subtract it in the image plane; this is described in the Image Analysis section of this cookbook. Here, we describe how to do continuum subtraction in the uv -plane.

The uv -plane continuum subtraction is performed by the `uvcontsub` task. First, determine which channels in your data cube do not have line emission, perhaps by forming a preliminary image as described in the next chapter. This image will also help you decide whether or not you need to come back and do uv -plane continuum subtraction at all.

The inputs to `uvcontsub` are:

```
# uvcontsub :: Continuum fitting and subtraction in the uv plane
vis                =      ''          # Name of input MS. Output goes to vis + ".contsub"
                                     # (will be overwritten if already exists)
field              =      ''          # Select field(s) using id(s) or name(s)
fitspw             =      ''          # Spectral window:channel selection for fitting the
                                     # continuum
combine            =      ''          # Data axes to combine for the continuum estimation
                                     # (none, or spw and/or scan)
solint             =      'int'       # Continuum fit timescale (int recommended!)
fitorder           =      0           # Polynomial order for the fits
spw                =      ''          # Spectral window selection for output
want_cont          =      False       # Create vis + ".cont" to hold the continuum estimate.
```

For each baseline, and over the timescale specified in `solint`, `uvcontsub` will provide a polynomial fit to the real and imaginary parts of the (continuum-only) channels specified in `fitspw` (using the standard `spw` selection syntax), and then subtract this model from all channels specified in `spw`, or from all channels in spectral windows of `fitspw` if `spw=''`. By setting the subparameter

`excludechannels=True`, the channel selection in `fitspw` will be inverted. In that case one can select the line channels themselves and/or corrupted channels that are not used in the continuum fit to the data. `fitspw` can also take frequency ranges, e.g.

```
fitspw='*:113.767~114.528GHz;114.744~115.447GHz'
```

where `*` indicates to go across all spws.

Typically, low orders for the polynomial work best, like 0th (a constant), or 1st order (a linear fit). Use higher orders with caution and check your results carefully.

Usually, one should set `solint='int'` which does no averaging and fits each integration. However, if the continuum emission comes from a small region around the phase center and `fitorder = 0`, then you can set `solint` larger (as long as it is shorter than the timescale for changes in the visibility function of the continuum). If your scans are short enough you can also use scan averaging with `combine='scan'` and `solint='inf'`. Be warned, setting `solint` too large will introduce “time smearing” in the estimated continuum and thus not properly subtract emission not at the phase center. Increasing `solint` speeds up the calculation but it does not improve the overall result quality of `uvcontsub` - although the continuum estimates of each baseline may be noisy (just like each visibility in a continuum MS may be noisy), it is better to use the ensemble of individual fits than to average the ensemble before fitting. Note that `plotms` can do time and baseline averaging on the fly to help you examine noisy data.

So, the recommended procedure is as follows:

- Finish calibration as described in the previous chapter.
- Use the `invert` or `clean` task on the split result to form an exploratory image that is useful for determining the line-free channels.
- Use `uvcontsub` with as low fit orders as possible to estimate and subtract the continuum from `vis`, and write the continuum-subtracted dataset to `vis + '.contsub'`.
- Use `clean` with `vis + '.contsub'` to make an image cube of the line emission.
- If a continuum image is desired, clean the line-free channels of the original MS with `mode='mfs'` and `spw=fitspw`. Note that using the line free channels directly is preferred over the imaging the 'continuum' model fitted by `uvcontsub`. The fitting procedure will also fit noise and artifacts which produce a nice line cube when subtracted, but the model may not represent the true underlying continuum.

For example, we perform *uv*-plane continuum subtraction on our NGC5921 dataset:

```
# Want to use channels 4-6 and 50-59 for continuum
uvcontsub(vis='ngc5921.usecase.ms',
field='N5921',
spw='', # all spw (only 0 in this data)
fitspw='0:4~7;50~59' # channels 4-6 and 50-59
```

```
solint='int',          # leave it at the default
fitorder=0)          # mean only
```

```
# You will see it made a new MS:
# ngc5921.usecase.ms.contsub"
```

4.7.7 Spectral regridding of the MS (cvel)

Although not strictly a calibration operation, spectral regridding of a MS is available to aid in calibration operations (e.g. continuum subtraction) and preparation for imaging. For this purpose, the `cvel` task has been developed.

The inputs are:

```
# cvel :: regrid an MS to a new spectral window / channel structure or frame
vis                =      ''          # Name of input measurement set
outputvis          =      ''          # Name of output measurement set
passall            =      False       # Pass through (write to output MS) non-selected data with
                                     # no change
field              =      ''          # Select field using field id(s) or field name(s)
spw                =      ''          # Select spectral window/channels
selectdata         =      True        # Other data selection parameters
  timerange        =      ''          # Range of time to select from data
  array            =      ''          # (sub)array indices
  antenna          =      ''          # Select data based on antenna/baseline
  scan             =      ''          # scan number range

mode               =      'channel'    # Regridding mode
  nchan            =      -1          # Number of channels in output spw (-1=all)
  start            =      0           # first input channel to use
  width            =      1           # Number of input channels to average
  interpolation     =      'linear'    # Spectral interpolation method

phasecenter        =      ''          # Image phase center: position or field index
restfreq           =      ''          # rest frequency (see help)
outframe           =      ''          # Output frame (not case-sensitive, ''=keep input frame)
veltype            =      'radio'     # velocity definition
hanning            =      False       # If true, Hanning smooth data before regridding to remove
                                     # Gibbs ringing.
```

The key parameters for the operation of `cvel` are the regridding mode, the output reference `outframe`, `veltype`, `restfreq` (which may be a list of rest frequencies to match the different spws) and the standard selection parameters (in particular `spw` and `field`).

The syntax for mode options (`'channel'`, `'velocity'`, `'frequency'`, `'channel.b'`) has been made compatible with the respective modes of `clean` (§ 5.2.5). The combination of selected `spw` and `mode` will determine the output channels and `spw(s)`:

```
spw = '0,1'; mode = 'channel'
```



```

# will produce a single spw containing all channels in spw 0 and 1
spw='0:5~28^2'; mode = 'channel'
# will produce a single spw made with channels (5,7,9,...,25,27)
spw = '0'; mode = 'channel': nchan=3; start=5; width=4
# will produce an spw with 3 output channels
# new channel 1 contains data from channels (5+6+7+8)
# new channel 2 contains data from channels (9+10+11+12)
# new channel 3 contains data from channels (13+14+15+16)
spw = '0:0~63^3'; mode='channel'; nchan=21; start = 0; width = 1
# will produce an spw with 21 channels
# new channel 1 contains data from channel 0
# new channel 2 contains data from channel 2
# new channel 21 contains data from channel 61
spw = '0:0~40^2'; mode = 'channel'; nchan = 3; start = 5; width = 4
# will produce an spw with three output channels
# new channel 1 contains channels (5,7)
# new channel 2 contains channels (13,15)
# new channel 3 contains channels (21,23)

```

The simplest use of `cvel` is to shift a single spectral window into an output frame without regridding. This is done with `mode='channel'`. For example:

```

cvel(vis='test_w3oh_nohann.ms',
     outputvis='test_w3oh_nohann_chanbary.ms',
     mode='channel',nchan=-1,start=0,width=1,
     interpolation='linear',
     phasecenter='',
     spw='',
     restfreq='1665.4018MHz',
     outframe='BARY')

```

does this for an observation of the OH line.

There is also a special `mode='channel_b'` that does not force a linear output frequency grid, e.g. for irregularly spaced/overlapping spectral windows), but is nominally faster. This is not equivalent to a `clean` output gridding mode, although `clean` will work on this spectral lattice.

Mode `channel` is intended to not interpolate between channels. It will perform binning if needed. For most scientific applications we recommend using the `mode='velocity''` and `mode='frequency'` options, as it is easiest to determine what the resulting channelization will be. For example:

```

cvel(vis='test_w3oh_nohann.ms',
     outputvis='test_w3oh_nohann_cvellsrk.ms',
     mode='velocity',nchan=45,start='-35.0km/s',width='-0.55km/s',
     interpolation='linear',
     phasecenter='',
     spw='',
     restfreq='1665.4018MHz',
     outframe='LSRK')

```

```

cvel(vis='test_w3oh_nohann.ms',
     outputvis='test_w3oh_nohann_cvelbary.ms',
     mode='velocity',nchan=45,start='-35.0km/s',width='-0.55km/s',
     interpolation='linear',
     phasecenter='',
     spw='',
     restfreq='1665.4018MHz',
     outframe='BARY')

```

will transform a MS into the LSRK and BARYcenter frames respectively.

The sign of the `width` parameter determines whether the channels run along increasing or decreasing values of frequency or velocity (i.e. if the cube is reversed or not).

The intent of `cvel` regridding is to transform channel labels and the visibilities to a spectral reference frame which is appropriate for the science analysis, e.g. from TOPO to LSRK, e.g. to correct for Doppler shifts throughout the time of the observation. Naturally, this will change the shape of the spectral features to some extent. According to the Nyquist theorem you should oversample a spectrum with twice the numbers of channels to retain the shape. Based on some tests, however, we recommend to observe with at least 3-4 times the number of channels for each significant spectral feature (like 3-4 channels per linewidth). This will minimize regridding artifacts in `cvel`.

If `cvel` has already established the grid that is desired for the imaging, `clean` should be run with the default channel mode (`> width=1`) or with exactly the same frequency/velocity parameters as was used in `cvel`. This will avoid additional regridding in `clean`. Hanning smoothing is optionally offered in `cvel`, but tests have shown that already the regridding process itself, if it involved a transformation from TOPO to a non-terrestrial reference frame, implies some smoothing (due to channel interpolation) such that Hanning smoothing may not be necessary.

The interpolation method `fftshift` calculates the transformed visibilities by applying a FFT, then a phase ramp, and then an inverse FFT. It will also perform pre-averaging, if necessary (this will increase the S/N). Note that if you want to use this interpolation method, your frequency grid needs to be equidistant, i.e. it only works in mode `velocity` with `veltype=radio`, in mode `frequency`, and in mode `channel` (in the latter only if the input grid is itself equidistant in frequency). Note also that, as opposed to all other interpolation methods, this method will apply a *constant* (frequency-independent) shift in frequency which is not fully correct in the case of large fractional bandwidth of the given spectral window.

4.7.8 UV-Plane Model Fitting (`uvmodelfit`)

It is often desirable to fit simple analytic source component models directly to visibility data. Such fitting has its origins in early interferometry, especially VLBI, where arrays consisted of only a few antennas and the calibration and deconvolution problems were poorly constrained. These methods overcame the calibration uncertainties by fitting the models to calibration-independent closure quantities and the deconvolution problem by drastically limiting the number of free parameters required to describe the visibilities. Today, even with larger and better calibrated arrays, it is still

desirable to use visibility model fitting in order to extract geometric properties such as the positions and sizes of discrete components in radio sources. Fits for physically meaningful component shapes such as disks, rings, and optically thin spheres, though idealized, enable connecting source geometry directly to the physics of the emission regions.

Visibility model fitting is carried out by the `uvmodelfit` task. The inputs are:

```
# uvmodelfit :: Fit a single component source model to the uv data:

vis      =      '' # Name of input visibility file
field    =      '' # field name or index
spw      =      '' # spectral window
selectdata = False # Activate data selection details
niter    =      5 # Number of fitting iterations to execute
comptype =      'P' # Component type (P=pt source,G=ell. gauss,D=ell. disk)
sourcepar = [1, 0, 0] # Starting guess (flux,xoff,yoff,bmajaxrat,bpa)
varypar  =      [] # Which parameters can vary in fit
outfile  =      '' # Optional output component list table
```

ALERT: This task currently only fits a single component.

The user specifies the number of non-linear solution iterations (`niter`), the component type (`comptype`), an initial guess for the component parameters (`sourcepar`), and optionally, a vector of Booleans selecting which component parameters should be allowed to vary (`varypar`), and a filename in which to store a CASA componentlist for use in other applications (`file`). Allowed `comptypes` are currently point 'P' or Gaussian 'G'.

The function returns a vector containing the resulting parameter list. This vector can be edited at the command line, and specified as input (`sourcepar`) for another round of fitting.

The `sourcepar` parameter is currently the only way to specify the starting parameters for the fit. For points, there are three parameters: I (total flux density), and relative direction (RA, Dec) offsets (in arcsec) from the observation's phase center. For Gaussians, there are three additional parameters: the Gaussian's semi-major axis width (arcsec), the aspect ratio, and position angle (degrees). It should be understood that the quality of the result is very sensitive to the starting parameters provided by the user. If this first guess is not sufficiently close to the global χ^2 minimum, the algorithm will happily converge to an incorrect local minimum. In fact, the χ^2 surface, as a function of the component's relative direction parameters, has a shape very much like the inverse of the absolute value of the dirty image of the field. Any peak in this image (positive or negative) corresponds to a local χ^2 minimum that could conceivably capture the fit. It is the user's responsibility to ensure that the correct minimum does the capturing.

Currently, `uvmodelfit` relies on the likelihood that the source is very near the phase center (within a beamwidth) and/or the user's savvy in specifying the starting parameters. This fairly serious constraint will soon be relieved somewhat by enabling a rudimentary form of uv-plane weighting to increase the likelihood that the starting guess is on a slope in the correct χ^2 valley.

Improvements in the works for visibility model fitting include:

- User-specifiable uv-plane weighting

- Additional component shapes, including elliptical disks, rings, and optically thin spheroids.
- Optional calibration pre-application
- Multiple components. The handling of more than one component depends mostly on efficient means of managing the list itself (not easy in command line options), which are currently under development.
- Combined component and calibration fitting.

Example (see Figure 4.11):

```
#
# Note: It's best to channel average the data if many channels
# before running a modelfit
#
split('ngc5921.ms', '1445+099_avg.ms',
      datacolumn='corrected', field='1445*', width='63')

# Initial guess is that it's close to the phase center
# and has a flux of 2.0 (a priori we know it's 2.47)
uvmodelfit('1445+099_avg.ms',      # use averaged data
           niter=5,                  # Do 5 iterations
           comptype='P',             # P=Point source, G=Gaussian, D=Disk
           sourcepar=[2.0,.1,.1],   # Source parameters for a point source
           spw='0',                 #
           outfile='gcal.cl')       # Output component list file

# Output looks like:
There are 19656 - 3 = 19653 degrees of freedom.
iter=0:  reduced chi2=0.0418509:  I=2,  dir=[0.1, 0.1] arcsec
iter=1:  reduced chi2=0.003382:   I=2.48562,  dir=[-0.020069, -0.0268826] arcsec
iter=2:  reduced chi2=0.00338012: I=2.48614,  dir=[0.00323428, -0.00232235] arcsec
iter=3:  reduced chi2=0.00338012: I=2.48614,  dir=[0.00325324, -0.00228963] arcsec
iter=4:  reduced chi2=0.00338012: I=2.48614,  dir=[0.00325324, -0.00228963] arcsec
iter=5:  reduced chi2=0.00338012: I=2.48614,  dir=[0.00325324, -0.00228963] arcsec
If data weights are arbitrarily scaled, the following formal errors
will be underestimated by at least a factor sqrt(reduced chi2). If
the fit is systematically poor, the errors are much worse.
I = 2.48614 +/- 0.0176859
x = 0.00325324 +/- 0.163019 arcsec
y = -0.00228963 +/- 0.174458 arcsec
Writing componentlist to file: /home/sandrock/smyers/Testing/Patch2/N5921/gcal.cl

# Fourier transform the component list to a model of the MS
ft('1445+099_avg.ms', complist='gcal.cl')

# Plot data versus uv distance
plotxy('1445+099_avg.ms', xaxis='uvdist', datacolumn='corrected')
```

```
# Specify green circles for model data (overplotted)
plotxy('1445+099_avg.ms', xaxis='uvdist', datacolumn='model',
       overplot=True, plotsymbol='go')
```

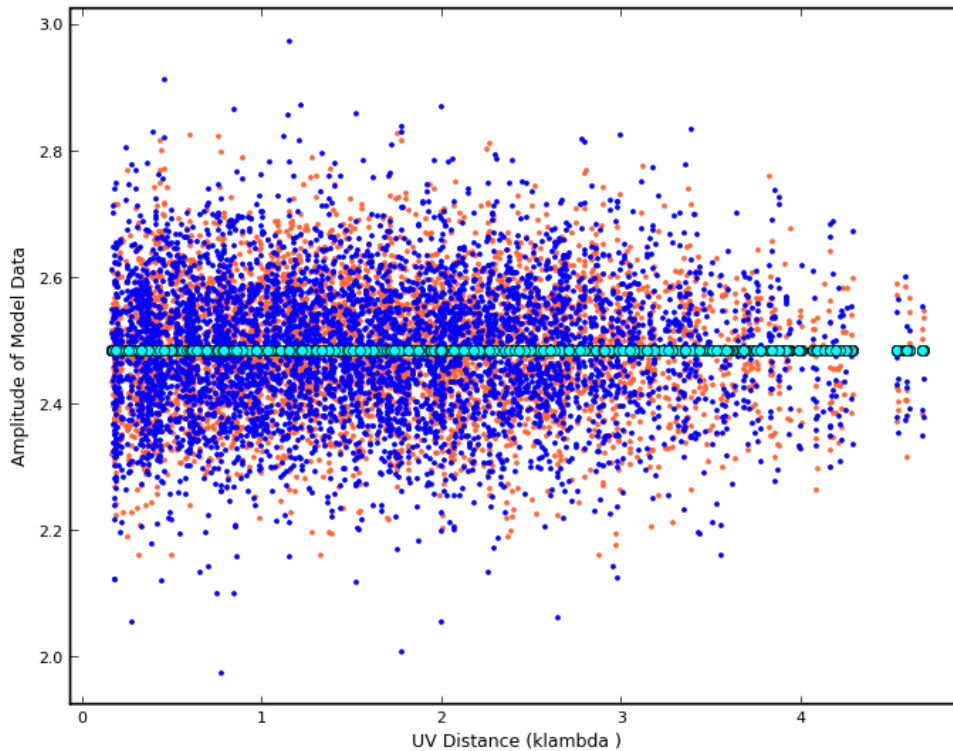


Figure 4.11: Use of `plotxy` to display corrected data (red and blue points) and uv model fit data (green circles).

4.7.9 Reweighting visibilities based on their scatter (`statwt`)

Alert: `statwt` is still an experimental task. Please check the results carefully and report any problems to the NRAO CASA helpdesk.

In most cases, the data that comes from the telescopes have the correct absolute or relative weights associated (absolute weights will be supplied once the VLA switched power application becomes standard; for ALMA the `Tsys` application is already in place). However, there are data sets where one would like to adjust the weights based on the scatter of the visibilities (typically as a function of time, antenna, and/or baseline). This calculation is performed by the task `statwt` that updates the `WEIGHT` and `SIGMA` columns of the Measurement Set. `statwt` inputs are:

```

# statwt :: Reweight visibilities according to their scatter
vis          =      ''          # Name of measurement set
dorms        =      False       # Use rms instead of stddev?
byantenna    =      False       # Estimate the noise per antenna -not
                                     # implemented (vs. per baseline)
fitspw       =      ''          # The signal-free spectral window:channels
                                     # to estimate the scatter from
fitcorr      =      ''          # The signal-free correlation(s) to estimate
                                     # the scatter from (not implemented)
combine      =      ''          # Let estimates span changes in spw, corr,
                                     # scan and/or state
timebin      =      '0s'        # Bin length for estimates (not implemented)
minsamp      =      2           # Minimum number of unflagged visibilities
                                     # for estimating the scatter
field        =      ''          # Select field using ID(s) or name(s)
spw          =      ''          # Select spectral window/channels
antenna      =      ''          # Select data based on antenna/baseline
timerange    =      ''          # Select data by time range
scan         =      ''          # Select data by scan numbers
intent       =      ''          # Select data by scan intents
array        =      ''          # Select (sub)array(s) by array ID number
correlation  =      ''          # Select correlations to reweight (DEPRECATED in CASA v4.5)
observation  =      ''          # Select by observation ID(s)
datacolumn   =      'corrected' # Which data column to calculate the scatter
                                     # from

```

`statwt` should only be run after all calibration steps have been performed. The parameter `dorms` switches from a scatter standard deviation to a root mean square scatter estimator. `datacolumn` specifies the column on which the task operates and the usual data selection parameters apply. Channels with strong RFI or a spectral line should be avoided for the calculation and good channel range should be specified via `fitspw`. In its current implementation, `statwt` uses data samples of an integration time interval but eventually wider sample intervals can be specified by the `timebin` parameter. Those samples are contained within a scan, `spw`, and polarization product but using the `combine` can relax this restriction. `minsamp` sets the minimum number of unflagged visibilities used for the calculation.

Alert: As of CASA 4.5, selection using `correlation` has been deprecated in `statwt`; in prior versions, this was not working correctly, and it is unlikely setting weights in a correlation-dependent manner is advisable.

4.7.10 Change the signs of visibility phases (conjugatevis)

`conjugatevis` is an easy task to flip the signs of the phases of visibilities, thus creating the complex conjugate numbers. The inputs are like:

```

# conjugatevis :: Change the sign of the phases in all visibility columns.
vis          =      ''          # Name of input visibility file.
spwlist      =      ''          # Spectral window selection
outputvis    =      ''          # Name of output visibility file
overwrite    =      False       # Overwrite the outputvis if it exists.

```

The task works on all scratch columns.

(This task was added early in JVLA commissioning, when some data suffered from a phase sign error.)

4.7.11 Manipulation of Ephemeris Objects

When an astronomical object has a proper motion, in particular objects in our solar system, a static (RA,Dec) position in the FIELD table of the MeasurementSet will not accurately describe the time-dependent position. Prior to CASA 4.2, there was no support for ephemeris objects other than the built-in reference frames for the Sun, the Moon, and the planets out to PLUTO. With CASA 4.2, several new features were introduced which help the user to attach an arbitrary ephemeris to a given field and work with the object from calibration to imaging.

4.7.11.1 Ephemeris tables

The CASA format for ephemeris tables has been defined was introduced in the early development stages of CASA in connection with the Measures module. The `me` tool permits position calculations based on ephemerides in this format. Two examples for such tables can be found in the distribution directory in subdirectory `data/ephemerides`: `VGEO` is an ephemeris of Venus in the geocentric reference frame while `VTOP` is an ephemeris for the same object in the `TOPO` reference frame for the observatory location of the VLA. With the introduction of solar system source models (Butler) in the `setjy` task, a nearly complete set of ephemerides for the larger bodies in our solar system had to be made available. These are stored in nearly the same format as the above examples `VGEO` and `VTOP` (but with a few enhancements) in directory `data/ephemerides/JPL-Horizons`. If your object's ephemeris is among those stored in `data/ephemerides/JPL-Horizons`, you can simply copy the ephemeris from there. Otherwise, you can request the ephemeris from the JPL-Horizons using the CASA commands (for example)

```
import recipes.ephemerides.request as jplreq
jplreq.request_from_JPL(objnam='Mars',startdate='2012-01-01',enddate='2013-12-31',
date_incr='0.1 d', get_axis_orientation=False,
get_axis_ang_orientation=True,
get_sub_long=True, use_apparent=False, get_sep=False,
return_address='YOUR_EMAIL_ADDRESS',
mailserver='YOUR_MAIL_SERVER_ADDRESS')
```

where you need to fill in the parameters `objnam`, `startdate`, `enddate`, `date_incr` (the time interval between individual ephemeris table entries), `return_address` (your email address where you want to receive the ephemeris), and `mailserver` (the smtp server through which you want to send the request email). The other parameters should be set as shown. Within a short time, you should receive the requested ephemeris as an email from NASA's JPL-Horizonssystem. Save the email into a file with the "save as" function of your mail client. See the next section on how to attach it to your dataset.

4.7.11.2 Using fixplanets to attach ephemerides to a field of a Measurement Set

In order to set the ephemeris of a given field in a Measurement Set, you can use the task `fixplanets` as in the following example:

```
fixplanets(vis='uid___A002_X1c6e54_X223.ms',
           field='Titan', fixuvw=True, direction='mytitanephemeris')
```

where you need to set the parameters `vis` to the name of your MS and the parameter `field` to the name of the field to which you want to attach the ephemeris. The parameter `direction` must be set to the name of your ephemeris table. Accepted formats are (a) the CASA format (as in `VGEO` or the ephemerides in `data/ephemerides/JPL-Horizons` as described above) and (b) the JPL-Horizons mail format which you obtain by saving an ephemeris email you received from JPL-Horizons. The parameter `fixuvw` should be set to `True` in order to trigger a recalculation of the UVW coordinates in your MS based on the new ephemeris. The task `fixplanets` can also be used for other field direction modifications. Please refer to the help text of the task.

Note that among the ephemerides in the directory `data/ephemerides/JPL-Horizons/` you should only use those ending in `'_J2000.tab'`. They are the ones in J2000 coordinates.

4.7.11.3 Use of the ephemeris after attachment

Once you have attached the ephemeris to a field of an MS, it will automatically be handled in tasks like `split` and `concat` which need to hand on the ephemeris to their output MSs. In particular `concat` recognizes when fields of the MSs to be concatenated use the same ephemeris and merges these fields **if** the time coverage of the provided ephemeris in the first MS also covers the observation time of the second MS. The ephemeris of the field in the first MS will then be used for the merged field. In order to inspect the ephemeris attached to a field, you can find it inside the `FIELD` subdirectory of your MS. The optional column `EPHEMERIS_ID` in the `FIELD` table points to the running number of the ephemeris table. A value of `-1` indicates that no ephemeris is attached. Note that in case an ephemeris is attached to a field, the `direction` column entries for that field in the `FIELD` table will be interpreted as an **offset** to the ephemeris direction and are therefore set to `(0.,0.)` by default. This offset feature is used in mosaic observations where several fields share the same ephemeris with different offsets. The `TIME` column in the `FIELD` table should be set to the beginning of the observation for that field and serves as the nominal time for ephemeris queries.

4.7.11.4 Spectral frame transformation to the rest frame of the ephemeris object in task `cvel`

The ephemerides contain radial velocity information. The task `cvel` can be used to transform spectral windows into the rest frame of the ephemeris object by setting the parameter `outframe` to `"SOURCE"` as in the following example:


```
cvel(vis='europa.ms',  
outputvis='cvel_europa.ms', outframe='SOURCE', mode = 'velocity',  
width = '0.3km/s', restfreq = '354.50547GHz')
```

This will make `cvel` perform a transformation to the GEO reference frame followed by an additional Doppler correction for the radial velocity given by the ephemeris for the each field. (Typically, this should happen after calibration and after splitting out the spectral widows and the target of interest). The result is an MS with a single combined spectral window in reference frame REST. From this frame, further transformations to other reference frames are not possible.

4.7.11.5 Ephemerides in ALMA datasets

The ALMA Science Data Model (the raw data format for ALMA data) now foresees an ephemeris table. However, this feature is not yet used in ALMA Cycle 1. Once this table will be filled at the observatory, the task `importasdm` will automatically translate it into an ephemeris table in CASA format and attach it to the respective fields.

4.8 Examples of Calibration

The data reduction tutorials on casaguides.nrao.edu provide walkthroughs for high and low frequency, spectral line and polarization calibration techniques.

Chapter 5

Synthesis Imaging

This chapter describes how to make and deconvolve images starting from calibrated interferometric data, possibly supplemented with single-dish data or an image made from single-dish data. This data must be available in CASA (see § 2 on importing data). See § 4 for information on calibrating synthesis data. In the following sections, the user will learn how to make various types of images from synthesis data, reconstruct images of the sky using the available deconvolution techniques, include single-dish information in the imaging process, and to prepare to use the results of imaging for improvement of the calibration process (“self-calibration”).

Inside the Toolkit:

The `im` tool handles synthesis imaging operations.

5.1 Imaging Tasks Overview

The current imaging and deconvolution tasks are:

- `clean` — calculate a deconvolved image with a selected clean algorithm, including mosaicing, or make a dirty image (§ 5.3),
- `feather` — combine a single dish and synthesis image in the Fourier plane (§ 5.6),
- `deconvolve` — image-plane only deconvolution based on the dirty image and beam, using one of several algorithms (§ 5.9),
- `pclean` — an experimental task for `clean` to work in a parallelized way for multi-node and core computing systems (§ 5.11)

There are also tasks that help you set up the imaging or interface imaging with calibration:

- `boxit` - create “cleanbox” deconvolution regions automatically from an image (§ 5.7.1),

- `ft` - add a source model to the MS (§ 5.8).

The full “tool kit” that allows expert-level imaging must still be used if you do not find enough functionality within the tasks above.

Information on other useful tasks and parameter setting can be found in:

- `listobs` — list what’s in a MS (§ 2.2.7),
- `split`— Write out new MS containing calibrated data from a subset of the original MS (§ section:cal.split),
- `cvel` — regrid a spectral MS onto a new frequency channel system (§ 4.7.7),
- data selection — general data selection syntax (§ 2.3),
- `viewer` — image display including region statistics and image cube slice and profile capabilities (§ 7).

5.2 Common Imaging Task Parameters

We now describe some parameters that are common to the imaging tasks. These should behave the same way in any imaging task that they are found in. These are in alphabetical order.

Inside the Toolkit:

The `im.setimage` method is used to set many of the common image parameters. The `im.advise` method gives helpful advice for setting up for imaging.

ALERT: `clean` tries to use up to four cores on the computer that it is running on. If this is not desired, the environment variable `OMP_NUM_THREAD` can be set to a lower value.

5.2.1 Parameter `cell`

The `cell` parameter defines the pixel size in the x and y axes for the output image. If given as floats or integers, this is the cell size in arc seconds, e.g.

```
cell=[0.5,0.5]
```

make 0.5'' pixels. You can also give the cell size in *quantities*, e.g.

```
cell=['1arcmin', '1arcmin']
```

If a single value is given, then square pixels of that size are assumed.

5.2.2 Parameter field

The `field` parameter selects the field indexes or names to be used in imaging. Unless you are making a mosaic, this is usually a single index or name:

```
field = '0'           # First field (index 0)
field = '1331+305'   # 3c286
field = '*'          # all fields in dataset
```

The syntax for `field` selection is given in § 2.3.2.

5.2.3 Parameter imagename

The value of the `imagename` parameter is used as the root name of the output image. Depending on the particular task and the options chosen, one or more images with names built from that root will be created. For example, the `clean` task run with `imagename='ngc5921'` a series of output images will be created with the names `ngc5921.clean`, `ngc5921.residual`, `ngc5921.model`, etc.

If an image with that name already exists, it will in general be overwritten. Beware using names of existing images however. If the `clean` is run using an `imagename` where `<imagename>.residual` and `<imagename>.model` already exist then `clean` will continue starting from these (effectively restarting from the end of the previous `clean`). Thus, if multiple runs of `clean` are run consecutively with the same `imagename`, then the cleaning is incremental (as in the `difmap` package).

The output image may also have a different beam per plane. For datasets with very large fractional bandwidth, `clean` will use a different PSF for each channel when the PSF changes by more than half a pixel as a function of frequency. To smooth to a common resolution, one can either use the parameter `resmooth` (§ 5.2.6) to smooth to the smallest common possible beam, `restoringbeam` for an arbitrary, larger beam, (§ 5.3.11), or the task `imsmooth` (§ 6.17) after `cleaning`. Data analysis tasks such as `immoments` in CASA support changing beams per plane.

5.2.4 Parameter imsize

The image size in numbers of pixels on the x and y axes is set by `imsize`. For example,

```
imsize = [640, 640]
```

makes a square image 640 pixels on a side. If a single value is given, then a square image of that dimension is made. The underlying algorithms work best for certain image sizes. If you pick a size where that algorithm will be particularly slow, the logger will send a warning message, suggesting the nearest optimal values. In general, the best performance is obtained with image sizes that are even and factorizable to 2,3,5,7 only. An easy rule of thumb would be $2^n \times 10$ where n is an integer number, like 160, 320, 640, 1280, 2560, etc.

5.2.5 Parameter mode

The mode parameter defines how the frequency channels in the synthesis MS are mapped onto the image. The allowed values are: `mfs`, `channel`, `velocity`, `frequency`. The mode parameter is expandable, with some options uncovering a number of sub-parameters, depending upon its value.

5.2.5.1 Mode mfs

```

mode           =      'mfs'      # Spectral gridding type (mfs, channel,
                                # velocity, frequency)
    nterms     =          1      # Number of terms used to model the sky
                                # frequency dependence (Note: nterms>1
                                # is under development)
    reffreq    =          ''      # Reference frequency for MFS (relevant
                                # only if nterms > 1),'' defaults to
                                # central data-frequency

```

The default `mode='mfs'` emulates multi-frequency synthesis in that each visibility-channel datum k with baseline vector \mathbf{B}_k at wavelength λ_k is gridded into the uv-plane at $\mathbf{u}_k = \mathbf{B}_k/\lambda_k$. The result is one or more images (depending on `nterms`), regardless of how many channels are in the input dataset. The first image plane is at the frequency given by the midpoint between the highest and lowest frequency channels in the input `spw(s)`. Currently, there is no way to choose the center frequency of the output image plane independently.

WideBand imaging (`mfs` with `nterms > 1`) is now available in CASA. This algorithm models the wide-band sky brightness as a linear combination of Gaussian-like functions whose amplitudes follow a Taylor-polynomial in frequency. The output images are a set of Taylor-coefficient images, from which spectral index and curvature maps are derived. The `reffreq` parameter sets the reference frequency ν_0 about which the Taylor expansion is done. The Taylor expansion is a polynomial in frequency:

$$I_{\nu}^{\text{sky}} = \sum_t I_t^{\text{sky}} \left(\frac{\nu - \nu_0}{\nu_0} \right)^t \quad (5.1)$$

I_t^{sky} an image of the t^{th} coefficient of the Taylor-polynomial expansion.

When Eq. 5.1 is applied on a source with a spectral index

$$I_{\nu}^{\text{sky}} = I_{\nu_0}^{\text{sky}} \left(\frac{\nu}{\nu_0} \right)^{\alpha + \beta \log(\nu/\nu_0)} \quad (5.2)$$

The Taylor terms I_t^{sky} can be used to constrain the sky brightness, α , and β through

$$I_{\nu_0}^{\text{sky}} = I_0^{\text{sky}} \quad (5.3)$$

$$\alpha = \frac{I_1^{\text{sky}}}{I_{\nu_0}^{\text{sky}}} = \frac{I_1^{\text{sky}}}{I_0^{\text{sky}}} \quad (5.4)$$

$$\beta = \frac{I_2^{\text{sky}}}{I_{\nu_0}^{\text{sky}}} - \frac{\alpha(\alpha - 1)}{2} = \frac{I_2^{\text{sky}}}{I_0^{\text{sky}}} - \frac{\alpha(\alpha - 1)}{2} \quad (5.5)$$

For more information, please see Rau, U. & Cornwell, T. J. 2011, “A multi-scale multi-frequency deconvolution algorithm for synthesis imaging in radio interferometry”, A&A, 532, 71

Alert: The MS-MFS (multiscale-multifrequency) algorithm in the current release is new and is still being developed/tested/debugged. Its basic operation has been tested on wide-band JVLA data for Stokes I imaging.

Explanation of the Parameters:

nterms: The number of terms in the Taylor polynomial used to model the frequency structure. **nterms**> 1 triggers MS-MFS. **nterms**= 1 triggers standard point-source clean or multi-scale-clean. Note: The choice of **nterms** follows the same rules used while fitting a polynomial to a 1D set of noisy data points. To prevent overfitting, the order of the polynomial needs to depend on the available signal-to-noise in the data. A very rough rule-of-thumb is as follows: For high SNR data (single channel SNR>100), and fields dominated by point-sources with spectral indices around -1.0 across a 2:1 bandwidth, choose **nterms**= 3 or 4. For lower SNR data (5 <SNR< 100), flatter spectra, or when there is significant extended emission, **nterms**= 2 is a much safer option. For very low SNR data (SNR< 5), choose **nterms**= 1).

reffreq: The reference frequency used to compute Taylor functions $[(\text{freq} - \text{reffreq})/(\text{reffreq})]^i$. If left blank (**reffreq**=”), it defaults to the middle frequency of the selected data. Note : For the current release, the use of **reffreq**=” is recommended.

multiscale: The MS-MFS algorithm always uses scale sizes set via the **multiscale** parameter. For point-source deconvolution, set **multiscale**=[0] (also the default). Note: Unlike standard ms-clean (**multiscale** = [0,6,10,...] with **nterms**=1), with higher **nterms** the largest specified scale size must lie within the sampled range of the interferometer. If not, there can be an ambiguity in the spectral reconstruction at very large spatial scales.

gridmode: Wideband W-Projection is supported, and can be triggered via **gridmode**=’widefield’.

modelimage: Supply a list of Taylor-coefficient images, to start the deconvolution from. If only one image is specified, it will be used as the model for the ’tt0’ image.

Output images: [xxx.image.tt0, xxx.image.tt1,...] : Images of Taylor coefficients that describe the frequency-structure. The ”tt0” image is the total-intensity image at the reference frequency, and is equivalent to ”xxx.image” obtained via standard imaging.

[xxx.image.alpha, xxx.image.beta] : Spectral index and spectral curvature at the reference-frequency. These are computed from tt0, tt1, tt2 only for regions of the image where there is sufficient signal-to-noise to compute them. These regions are chosen via a threshold on the intensity image (tt0) computed as MAX(userthreshold*5 , peakresidual/10). This threshold is reported in the logger. Elsewhere, the values are currently set to zero.

[xxx.image.alpha.error] contains the errors of the spectral index solutions.

The following is a list of differences between MS-MFS (`nterms > 1`) and standard imaging, in the current CASA release.

1. Iterations always proceed as cs-clean major/minor cycles, and uses the full psf during minor cycle iterations. There are currently no user-controls on the `cyclespeedup`, and the flux-limit per major cycle is chosen as 10% of the peak residual. In future releases, this will be made more adaptive/controllable.
2. Currently, the following options are not supported for `nterms > 1`: `psfmode`, `pbcorr`, `minpb`, `imagermode='mosaic'`, `gridmode='projection'`, `cyclespeedup`, and allowed are one of Stokes I, Q, U, V, RR, LL, XX, YY at a time. More options and combinations are currently under development and testing. Under 'Using CASA' → 'Other Documentation' → 'Imaging Algorithms in CASA' you can find the latest implementations.

5.2.5.2 Mode channel

ALERT: Note that `mode='channel'` is intended as a shortcut to produce a cube based on the input MS channelization. It will be in the frame of the input MS. We recommend that users instead use the `'velocity'` and `'frequency'` modes which will produce cubes in other frames with more control of the cube spacing. These modes have defaults that will work from the MS spacing, reproducing the action of `mode='channel'`.

If `mode='channel'` is chosen, then an image cube will be created. This is an expandable parameter, with dependent parameters:

```
mode           = 'channel' # Spectral image definition(mfs,
                          # channel, velocity,frequency)
  nchan        =      -1 # Number of channels (planes) in output image
  start        =         0 # first input channel to use
  width        =         1 # Number of input channels to average
  interpolation = 'nearest' # Spectral interpolation(nearest, linear, cubic)
```

The default `nchan=-1` will automatically produce a cube with the number of channels needed to span the (regridded) spectral windows of the MS. If multiple MSs are used, the spectral frames of these need to be identical, e.g. LSRK¹. **ALERT:** This often results in extra blank channels at the beginning and end of the image cube, so it is usually more precise to specify `nchan` and `start` to get what you want. For best results, we also recommend `'nearest'` interpolation for the `mode=channel`.

The channelization of the resulting image is determined by the channelization in the MS of `vis` of the first `spw` specified (the “reference `spw`”). The actual channels to be gridded and used in

¹Note that when TOPO is used, it refers to a time stamp at a given observation date. If more than one observation in TOPO is specified, this may lead to vastly erroneous values. Any conversion from TOPO to other frames such as BARY and LSRK should be performed for each individual observation, prior to clean or concatenation

the clean are selected via the `spw` parameter as usual. The resulting image cube will have `nchan` channels spaced evenly in frequency. The first output channel will be located at the frequency of channel `start` in the (first) reference `spw` (independent of what channels are selected using `spw`). If `width > 1`, then input MS channels with centers within a frequency range given by $(\text{width} + 1)/2$ times the reference `spw` spacing will be gridded together (as in `mode = 'mfs'` above) into the channels of the output image cube. The output channel spacing is thus given by `width` channels in the reference `spw` of the MS.

The `interpolation` sub-parameter (§ 5.2.5.5) sets how channels are gridded into the image cube planes. For `'nearest'`, the channels in `spw` beyond the first are mapped into the nearest output image channel within half a channel (if any). Otherwise, the chosen interpolation scheme will be used. Image channels that lie outside the MS frequency range or have no data mapped to them will be blank in the output image, but will be in the cube.

Example:

```
mode      = 'channel'
nchan     =      46
start     =       5
width     =       1
```

which will produce a 46-channel cube starting with channel 5 of the MS with the same channel width as the MS. *Note: the `start` channel is in reference to the channels in the MS, not the subset selected by `spw`.*

5.2.5.3 Mode frequency

For `mode='frequency'`, an output image cube is created with `nchan` channels spaced evenly in frequency.

```
mode      = 'frequency' # Spectral image definition(mfs,
                        # channel, velocity,frequency)
nchan     =      -1     # Number of channels (planes) in output image
start     =          '' # Frequency of first image channel:
                        # e.q. '1.4GHz'(''=default)
width     =          '' # Image channel frequency width:
                        # e.g '1.0kHz'(''=default)
interpolation = 'linear' # Spectral interpolation(nearest, linear, cubic)
outframe  =          '' # velocity frame of output image
```

The frequency of the first output channel is given by `start` and spacing by `width`. Output channels have width also given by `width`. The sign of `width` determines whether the output channels ascend or descend in frequency. Data from the input MS with centers that lie within one-half an input channel overlap of the frequency range of $\pm \text{width}/2$ centered on the output channels are gridded together.

The defaults are designed to safely choose output cube channels to span the input MS(s). The default `nchan=-1` will choose the number of channels needed to span the frequencies of the channels

in the MS. The defaults `start=''` and `width=''` will use the channel frequency and width of the first channel of the first specified spectral window selected in `spw`. **ALERT:** As in “channel” mode, this is currently the first channel of the first `spw`, not the first channel selected from that `spw`.

The `interpolation` sub-parameter (§ 5.2.5.5) sets how channels are gridded into the image cube planes.

Using the NGC5921 dataset as an example:

```
mode          = 'frequency'
nchan         =      21
start        = '1412.830MHz'
width        =    '50kHz'
outframe     =    'LSRK'
```

would produce a 21-channel output cube with 50 kHz wide channels rather than the default channelization of the MS (24.4 kHz).

5.2.5.4 Mode velocity

If `mode='velocity'` is chosen, then an output image cube with `nchan` channels will be created, with channels spaced evenly in velocity. Parameters are:

```
mode          = 'velocity' # Spectral image definition(mfs,
                           # channel, velocity,frequency)
nchan         =      -1 # Number of channels (planes) in output image
start         =      '' # Velocity of first image channel:
                       # e.g '0.0km/s'(''=default)
width         =      '' # Image channel velocity width: e.g
                       # '-1.0km/s'(''=default)
interpolation = 'linear' # Spectral interpolation(nearest,
                           # linear, cubic)
outframe      =      '' # velocity reference frame of output
                       # image; '' =input
veltype       =    'radio' # velocity definition
```

Note that velocities are calculated with respect to the rest frequency in the MS or specified through the `restfreq` parameter (§ 5.2.8).

The velocity of the first output channel is given by `start` and spacing by `width`. Averaging is as in `mode='frequency'`. The `interpolation` sub-parameter (§ 5.2.5.5) sets how channels are gridded into the image cube planes.

The defaults are designed to safely choose output cube channels to span the input MS(s). The default `nchan=-1` will choose the number of channels needed to span the velocities of the channels in the MS. The defaults `start=''` and `width=''` will use the channel velocity and width of the first channel of the first specified spectral window selected in `spw`. **ALERT:** As in “channel” mode, this is currently the first channel of the first `spw`, not the first channel selected from that `spw`.

Again, using the NGC5921 dataset as an example:

```

mode      = 'velocity'
  nchan   =      21
  start   = '1383.0km/s'
  width   = '10km/s'
  outframe = 'LSRK'

```

Note that in this case the velocity axis runs forward, as opposed to the default channelization for 'channel' or 'frequency'.

5.2.5.5 Sub-parameter interpolation

The `interpolation` sub-parameter controls how spectral channels in the MS are gridded into the output image cube. This is available in all modes except 'mfs'. The options are: 'nearest', 'linear', 'cubic'.

For 'nearest', the channels in `spw` beyond the first are mapped into the nearest output image channel within half a channel (if any).

For 'linear', the channels are gridded into the planes using weights given by a linear function of the frequency of the MS channel versus the plane. Each input channel will be mapped to 1 or 2 output planes. For most users, this is the best choice.

For 'cubic', the channels are gridded using a cubic interpolation function.

'Linear' and 'cubic' interpolation methods require that there are two datapoints that sandwich your new, regridded bin. This can introduce edge effects like in the first or last channel or adjacent to flagged channels where data is only available on one side of the spectrum. `interpolation='nearest'` will avoid such edge effects but may not work so well for data with `spws` that overlap. For `mode='velocity'` or 'frequency', 'linear' interpolation usually works best and for `mode='channel'` the 'nearest' interpolation method is superior. But this could be different for your dataset and you should carefully check your results.

5.2.6 Parameter resmooth

For large cubes, the `psf` will change as a function of frequency. `clean` will produce cubes with different synthesized beams per plane. All CASA analysis tasks can deal with such cubes. If one would like a common `psf` for all planes, typically the smallest possible beam, one can invoke the `resmooth` Boolean parameter. Alternatively, the cube can be convolved to the smallest common beam in a separate step vis `imsmooth` (see Sect. 6.17).

5.2.7 Parameter phasecenter

The `phasecenter` parameter indicates which of the field IDs should be used to define the phase center of the mosaic image, or what that phase center is in RA and Dec. The default action is to use the first one given in the `field` list.

For example:

```

phasecenter='5'                # field 5 in multi-src ms
phasecenter='J2000 19h30m00 -40d00m00' # specify position

```

Note that the format for angles prefers to use `hm` for RA/time units and `dm` for Dec/Angle units as separators. The colon `::` separator is interpreted as RA/time even if its used for the Dec, so be careful not to copy/paste from other sources.

5.2.8 Parameter `restfreq`

The value of the `restfreq` parameter, if set, will over-ride the rest frequency in the header of the first input MS to define the velocity frame of the output image.

ALERT: The `restfreq` parameter takes the options of transitions and frequencies as in the corresponding `plotxy` parameter (§ 3.3.2.12), but the frame information is controlled under the `mode` parameter (§ 5.2.5).

For example:

```
restfreq='115.2712GHz',
```

will set the rest frequency to that of the CO 1-0 line.

ALERT: Setting `restfreq` explicitly here in `clean` is good practice, and may be necessary if your MS has been concatenated from different files for different spectral windows (§ 2.2.12).

5.2.9 Parameter `spw`

The `spw` parameter selects the spectral windows that will be used to form the image, and possibly a subset of channels within these windows.

The `spw` parameter is a string with an integer, list of integers, or a range, e.g.

```

spw = '1'           # select spw 1
spw = '0,1,2,3'    # select spw 0,1,2,3
spw = '0~3'        # same thing using ranges

```

You can select channels in the same string with a `:` separator, for example

```

spw = '1:10~30'    # select channels 10-30 of spw 1
spw = '0:5~55,3:5;6;7' # chans 5-55 of spw 0 and 5,6,7 of spw 3

```

This uses the standard syntax for `spw` selection is given in § 2.3.3. See that section for more options.

Note that the order in which multiple `spws` are given is important for `mode = 'channel'`, as this defines the origin for the channelization of the resulting image.

5.2.10 Parameter stokes

The `stokes` parameter specifies the Stokes parameters for the resulting images. Note that forming Stokes Q and U images requires the presence of cross-hand polarizations (e.g. RL and LR for circularly polarized systems such as the VLA) in the data. Stokes V requires both parallel hands (RR and :LL) for circularly polarized systems or the cross-hands (XY and YX) for linearly polarized systems such as ALMA and ATCA.

This parameter is specified as a string of up to four letters and can indicate stokes parameters themselves, Right/Left hand polarization products, or linear polarization products (X/Y). For example,

```

stokes = 'I'           # Intensity only
stokes = 'IQU'        # Intensity and linear polarization
stokes = 'IV'         # Intensity and circular polarization
stokes = 'IQUV'       # All Stokes imaging
stokes = 'RR'         # Right hand polarization only
stokes = 'XXYY'       # Both linear polarizations

```

are common choices (see the inline help of `clean` for a full range of possible options). The output image will have planes (along the “polarization axis”) corresponding to the chosen Stokes parameters.

If as input to deconvolution tasks such as `clean`, the `stokes` parameter includes polarization planes other than I, then choosing `psfmode='hogbom'` (§ 5.3.1.2) or `psfmode='clarkstokes'` (§ 5.3.1.3) will clean (search for components) each plane sequentially, while `psfmode='clark'` (§ 5.3.1.1) will deconvolve jointly.

Alert: As of Release 3.2, `clean` expects that all input polarizations are present. E.g. if you have RR and LL dual polarization data and you flagged parts of RR but not LL, `clean` will ignore **both** polarizations in slice. It is possible to split out a polarization product with `split` and image separately. But you will not be able to combine these part-flagged data in the uv-domain. We will remove that restriction in a future CASA release.

5.2.11 Parameter uvtaper

This controls the radial weighting of visibilities in the uv-plane (see § 5.2.12 below) through the multiplication of the visibilities by the Fourier transform of an elliptical Gaussian. This is itself a Gaussian, and thus the visibilities are “tapered” with weights decreasing as a function of uv-radius.

The `uvtaper` parameter expands the menu upon setting `uvtaper=True` to reveal the following sub-parameters:

```

uvtaper      =      True   # Apply additional uv tapering of visibilities.
  outertaper =      []    # uv-taper on outer baselines in uv-plane
  innertaper =      []    # uv-taper in center of uv-plane (not
implemented)

```

The sub-parameters specify the size and optionally shape and orientation of this Gaussian in the uv-plane or optionally the sky plane. The `outertaper` refers to a Gaussian centered on the origin of the uv-plane.

Some examples:

```

outertaper=[]                # no outer taper applied
outertaper=['5klambda']     # circular uv taper FWHM=5 kilo-lambda
outertaper=['5klambda','3klambda','45.0deg'] # elliptical Gaussian
outertaper=['10arcsec']    # on-sky FWHM 10"
outertaper=['300.0']       # 300m in aperture plane

```

Note that if no units are given on the taper, then the default units are assumed to be meters in aperture plane.

ALERT: The `innertaper` option is not yet implemented.

5.2.12 Parameter weighting

In order to image your data, we must have a map from the visibilities to the image. Part of that map, which is effectively a convolution, is the weights by which each visibility is multiplied before gridding. The first factor in the weighting is the “noise” in that visibility, represented by the data weights in the MS (which is calibrated along with the visibility data). The weighting function can also depend upon the uv locus of that visibility (e.g. a “taper” to change resolution). This is actually controlled by the `uvtaper` parameter (see § 5.2.11). The weighting matrix also includes the convolution kernel that distributes that visibility onto the uv-plane during gridding before Fourier transforming to make the image of the sky. This depends upon the density of visibilities in the uv-plane (e.g. “natural”, “uniform”, “robust” weighting).

Inside the Toolkit:
 The `im.weight` method has more weighting options than available in the imaging tasks. See the **User Reference Manual** for more information on imaging weights.

The user has control over all of these.

ALERT: You can find a weighting description in the online User Reference Manual at:

<http://casa.nrao.edu/docs/casaref/imager.weight.html>

The `weighting` parameter expands the menu to include various sub-parameters depending upon the mode chosen:

5.2.12.1 ‘natural’ weighting

For `weighting='natural'`, visibilities are weighted only by the data weights, which are calculated during filling and calibration and should be equal to the inverse noise variance on that visibility.

Imaging weight w_i of sample i is given by

$$w_i = \omega_i = \frac{1}{\sigma_k^2} \quad (5.6)$$

where the data weight ω_i is determined from σ_i is the rms noise on visibility i . When data is gridded into the same uv-cell for imaging, the weights are summed, and thus a higher uv density results in higher imaging weights. No sub-parameters are linked to this mode choice. It is the default imaging weight mode, and it should produce “optimum” image with the lowest noise (highest signal-to-noise ratio). Note that this generally produces images with the poorest angular resolution, since the density of visibilities falls radially in the uv-plane

5.2.12.2 ‘uniform’ weighting

For `weighting = ‘uniform’`, the data weights are calculated as in ‘natural’ weighting. The data is then gridded to a number of cells in the uv-plane, and after all data is gridded the uv-cells are re-weighted to have “uniform” imaging weights. This pumps up the influence on the image of data with low weights (they are multiplied up to be the same as for the highest weighted data), which sharpens resolution and reduces the sidelobe level in the field-of-view, but increases the rms image noise. No sub-parameters are linked to this mode choice.

For uniform weighting, we first grid the inverse variance ω_i for all selected data onto a grid with uv cell-size given by $2/FOV$ where FOV is the specified field of view (defaults to the image field of view). This forms the gridded weights W_k . The weight of the i -th sample is then:

$$w_i = \frac{\omega_i}{W_k}. \quad (5.7)$$

5.2.12.3 ‘superuniform’ weighting

The `weighting = ‘superuniform’` mode is similar to the ‘uniform’ weighting mode but there is now an additional `npixels` sub-parameter that specifies a change to the number of cells on a side (with respect to uniform weighting) to define a uv-plane patch for the weighting renormalization. If `npixels=0` you get uniform weighting.

5.2.12.4 ‘radial’ weighting

The `weighting = ‘radial’` mode is a seldom-used option that increases the weight by the radius in the uv-plane, i.e.

$$w_i = \omega_i \cdot \sqrt{u_i^2 + v_i^2}. \quad (5.8)$$

Technically, I would call that an inverse uv-taper since it depends on uv-coordinates and not on the data per-se. Its effect is to reduce the rms sidelobes for an east-west synthesis array. This option has limited utility.

5.2.12.5 'briggs' weighting

The `weighting = 'briggs'` mode is an implementation of the flexible weighting scheme developed by Dan Briggs in his PhD thesis. See:

<http://www.aoc.nrao.edu/dissertations/dbriggs/>

This choice brings up the sub-parameters:

```
weighting      = 'briggs'  # Weighting to apply to visibilities
robust         = 0.0       # Briggs robustness parameter
npixels       = 0         # number of pixels to determine uv-cell size 0=> field of view
```

The actual weighting scheme used is:

$$w_i = \frac{\omega_i}{1 + W_k f^2} \quad (5.9)$$

where W_k is defined as in `uniform` and `superuniform` weighting, and

$$f^2 = \frac{(5 * 10^{-R})^2}{\frac{\sum_k W_k^2}{\sum_i \omega_i}} \quad (5.10)$$

and R is the robust parameter.

The key parameter is the `robust` parameter, which sets R in the Briggs equations. The scaling of R is such that $R = 0$ gives a good trade-off between resolution and sensitivity. The `robust` R takes value between -2.0 (close to uniform weighting) to 2.0 (close to natural).

Superuniform weighting can be combined with Briggs weighting using the `npixels` sub-parameter. This works as in `'superuniform'` weighting (§ 5.2.12.3).

5.2.12.6 'briggsabs' weighting

For `weighting='briggsabs'`, a slightly different Briggs weighting is used, with

$$w_i = \frac{\omega_i}{W_k R^2 + 2\sigma_R^2} \quad (5.11)$$

where R is the robust parameter and σ_R is the `noise` parameter.

This choice brings up the sub-parameters:

```
weighting      = 'briggsabs' # Weighting to apply to visibilities
robust         = 0.0       # Briggs robustness parameter
noise         = '0.0Jy'    # noise parameter for briggs weighting when rmode='abs'
npixels       = 0         # number of pixels to determine uv-cell size 0=> field of view
```

Otherwise, this works as `weighting='briggs'` above (§ 5.2.12.5).

5.2.13 Parameter `vis`

The value of the `vis` parameter is either the name of a single MS, or a list of strings containing the names of multiple MSs, that should be processed to produce the image. The MS referred to by the first name in the list (if more than one) is used to determine properties of the image such as channelization and rest frequency.

For example,

```
vis = 'ngc5921.ms'
```

set a single input MS, while

```
vis = ['ngc5921_day1.ms', 'ngc5921_day2.ms', 'ngc5921_day3.ms']
```

points to three separate measurement sets that will be gridded together to form the image. This means that you do not have to concatenate datasets, for example from different configurations, before imaging.

For the multiple MS case, all selection commands like `field`, `spw`, etc. are lists that refer to the list of input MSs, like

```
spw=['1:2~9', '0:10~22', '<2']
field=['0', 'ngc5921', '12']
```

will use the first entry of each selection criterion and apply it to the first dataset (`spw='1:2~9'` and `field='0'` to `'ngc5921_day1.ms'`), the second selection criterion to the second dataset etc.

5.2.14 Primary beams in imaging

The CASA imaging task and tools use primary beams based on models for each observatory's antenna types. In addition to different antenna diameters, different functions may be used.

The voltage patterns are based on the following antenna primary beams, based on the `TELESCOPE_NAME` keyword in the `OBSERVATION` table:

VLA — Airy disk fitted to measurement. Note that a R/L beam squint is also included with feed dependent angle;

ALMA — Airy disk for 12m dish with a blockage of 1m;

ATA — Airy disk for 6m dish;

ATCA — polynomial fitted to measurement of main lobe;

BIMA, HATCREEK — Gaussian with halfwidth of $\lambda/2D$;

CARMA — Airy patterns for the BIMA or OVRO dish sizes as appropriate;

GBT — polynomial fitted to measurement of main lobe;

GMRT — VLA Airy disk scaled to 45.0m;

IRAMPDB — Airy disk for dish of 15m with a blockage of 1m;

NRAO12M — VLA beam scaled to 12m;

OVRO — VLA Airy disk scaled to 10.4m;

SMA — Spheroidal function fit to FWHM;

WSRT — polynomial fitted to measurement of main lobe;

If the telescope name is unknown, or is CARMA or ALMA, then the `DISH_DIAMETER` in the `ANTENNA` table is used with a scaled VLA pattern.

In mosaicking mode, `clean` will use frequency-dependent primary beams. It also appears that Airy or spheroidal beams are best behaved for mosaics (see § 5.3.15).

5.3 Deconvolution using CLEAN (`clean`)

To create an image and then deconvolve it with the CLEAN algorithm, use the `clean` task. This task will work for single-field data, or for multi-field mosaics (§ 5.3.15), in both narrow and wide-field imaging modes.

ALERT: For large fractional bandwidths the `psf` in `clean` may vary considerably with frequency in data cubes. To accommodate this fact we have introduced a per-plane `psf` (dirty beam) when the change is larger than half the size of a pixel. Analysis tasks in CASA can deal with such beam variation. If a single beam size is requested, `imsmooth` can be invoked on the `clean` products to smooth to a common, uniform beam for all channels.

Toolkit Note: MEM is not included in `clean`, but is available in the toolkit.

`clean` will use the `CORRECTED_DATA` column from your measurement set if it exists. If that column is not available, it will use `DATA`. The `clean` task utilizes many of the common imaging parameters. These are described above in § 5.2. There are also a number of parameters specific to `clean`. These are listed and described below.

The default inputs to `clean` are:

```
# clean :: Deconvolve an image with selected algorithm
vis                =      '' # name of input visibility file
imagenam          =      [''] # Pre-name of output images
outlierfile       =      '' # Text file with image names, sizes, centers
field             =      '' # Field Name
spw               =      '' # Spectral windows:channels: '' is all
```

```

selectdata      =      False  # Other data selection parameters
mode            =      'mfs'  # Type of selection (mfs, channel, velocity,frequency)
      nterms    =      1      # Number of taylor terms to use
      #         # for modeling the sky frequency dependence
      reffreq   =      ''     # Reference frequency for MFS
      #         # (relevant only if nterms > 1)

gridmode       =      ''     # The kind gridding kernel to be
      #         # used for FFT-based transforms
niter          =      500    # Maximum number of iterations
gain           =      0.1    # Loop gain for cleaning
threshold      =      '0.0mJy' # Flux level to stop cleaning. Must include units
psfmode        =      'clark' # method of PSF calculation to use during minor cycles
imagermode     =      ''     # Use csclean or mosaic. If '', use psfmode
multiscale     =      []     # deconvolution scales (pixels);
      #         # [] = default standard clean

interactive     =      False  # use interactive clean (with GUI viewer)
mask           =      []     # cleanbox(es), mask image(s),
      #         # and/or region(s) used in cleaning
imsize         =      [256, 256] # x and y image size in pixels,
      #         # symmetric for single value
cell           =      ['1.0arcsec', '1.0arcsec'] # x and y cell size. default unit arcsec
phasecenter    =      ''     # Image phase center: position or field index
restfreq       =      ''     # rest frequency to assign to image (see help)
stokes         =      'I'    # Stokes params to image (eg I,IV, QU,IQUV)
weighting      =      'natural' # Weighting of uv (natural, uniform, briggs, ...)
uvtaper        =      False  # Apply additional uv tapering of visibilities.
modelimage     =      ''     # Name of model image(s) to initialize cleaning
restoringbeam  =      ['']   # Output Gaussian restoring beam for CLEAN image
pbcor          =      False  # Output primary beam-corrected image
minpb         =      0.1    # Minimum PB level to use
usescratch     =      False  # True if to save model
      #         # visibilities in MODEL_DATA column

```

The `clean` task will produce a number of output images based on the root name given in `imagename`. These include:

```

<imagename>.clean.image          # the restored image
<imagename>.clean.flux           # the effective response (e.g. for pbcor)
<imagename>.clean.flux.pbcovrage # the PB coverage (ftmachine='mosaic' only)
<imagename>.clean.model         # the model image
<imagename>.clean.residual      # the residual image
<imagename>.clean.psf           # the synthesized (dirty) beam

```

The `mode`, `psfmode`, `imagermode`, and `weighting` parameters open up other sub-parameters. These are detailed in the common imaging task parameters section (§ 5.2). The `gridmode` parameter (§ 5.3.13) is available to select more advanced imaging options such as widefield imaging and beam squint correction.

A typical setup for `clean` on the NGC5921 dataset, after setting parameter values, might look like:

```

vis            = 'ngc5921.usecase.ms.contsub' # Name of input visibility file

```

```

imagename      = 'ngc5921.usecase.clean' # Pre-name of output images
field          = '0' # Field Name
spw            = '' # Spectral windows:channels: '' is all
selectdata     = False # Other data selection parameters
mode           = 'channel' # Type of selection (mfs, channel, velocity, frequency)
    nchan       = 46 # Number of channels (planes) in output image
    start       = 5 # first input channel to use
    width       = 1 # Number of input channels to average
    interpolation = 'linear' # Spectral interpolation (nearest, linear, cubic)

gridmode       = '' # The kind gridding kernel to be used for
                # FFT-based transforms
niter          = 6000 # Maximum number of iterations
gain           = 0.1 # Loop gain for cleaning
threshold      = 8.0 # Flux level to stop cleaning. Must include units
psfmode        = 'clark' # method of PSF calculation to use during minor cycles
imagermode     = '' # Use csclean or mosaic, or image-plane only if ''
multiscale     = [] # set deconvolution scales (pixels)
interactive    = False # use interactive clean (with GUI viewer)
mask           = [108, 108, 148, 148] # cleanbox(es), mask image(s),
                # and/or region(s)

imsize         = [256, 256] # x and y image size in pixels
cell           = [15.0, 15.0] # x and y cell size. default unit arcsec
phasecenter    = '' # Image phase center: position or field index
restfreq       = '' # rest frequency to assign to image (see help)
stokes         = 'I' # Stokes params to image (eg I,IV, QU,IQUV)
weighting      = 'briggs' # Weighting to apply to visibilities
    robust      = 0.5 # Briggs robustness parameter
    npixels     = 0 # uv-cell size in pixels 0=> field of view

uvtaper        = False # Apply additional uv tapering of visibilities.
modelimage     = '' # Name of model image(s) to initialize cleaning
restoringbeam  = [''] # Output Gaussian restoring beam for CLEAN image
pbcor          = False # Output primary beam-corrected image
minpb          = 0.1 # Minimum PB level to use

```

An example of the `clean` task to create a continuum image from many channels is given below:

```

clean(vis='ggtau.1mm.split.ms', # Use data in ggtau.1mm.split.ms
      imagename='ggtau.1mm', # Name output images 'ggtau.1mm.*' on disk
      psfmode='clark', # Use the Clark CLEAN algorithm
      imagermode='', # Do not mosaic or use csclean
      mask='', # Do not use clean box or mask
      niter=500, gain=0.1, # Iterate 500 times using gain of 0.1
      mode='mfs', # multi-frequency synthesis (combine channels)
      spw='0~2:2~57', # Combine channels from 3 spectral windows
      field='0', #
      stokes='I', # Image stokes I polarization
      weighting='briggs', # Use Briggs robust weighting
      rmode='norm', robust=0.5, # with robustness parameter of 0.5
      cell=[0.1,0.1], # Using 0.1 arcsec pixels

```

```
imsize=[256,256])          # Set image size = 256x256 pixels
```

This example will clean the entire inner quarter of the primary beam. However, if you want to limit the region over which you allow the algorithm to find clean components then you can make a deconvolution region (or mask). To use a deconvolution region, box, or mask, set the `mask` parameter (§ 5.3.6).

For example, you can set up a simple 'cleanbox' region. To do this, make a first cut at the image and clean the inner quarter. Then use the `viewer` to look at the image and get an idea of where the emission is located. You can use the `viewer adjustment` panel to view the image in pixel coordinates and read out the pixel locations of your cursor.

Then, you can use those pixel read-outs you just go to define a clean box region with the CASA region format described in Chapter D. For example, say you have a continuum source near the center of your image between the pixel coordinates [80,80] and [120,120], you may use the rectangular region:

```
mask='box[[80pix,80pix],[120pix,120pix]]'
```

For more complicated and multiple clean regions, it will be best to use the `viewer` to create them interactively or to create a region file (Chapter D) and use that file as an input like:

```
mask='myregions.txt'
```

The following are the `clean` specific parameters and their allowed values, followed by a description of carrying out interactive cleaning.

5.3.1 Parameter `psfmode`

The `psfmode` parameter chooses the “algorithm” that will be used to calculate the synthesized beam for use during the minor cycles in the image plane. The value types are strings. Allowed choices are 'clark' (default) and 'hogbom'.

5.3.1.1 The clark algorithm

In the 'clark' algorithm, the cleaning is split into minor and major cycles. In the minor cycles only the brightest points are cleaned, using a subset of the point spread function. In the major cycle, the points thus found are subtracted correctly by using an FFT-based convolution. This algorithm is reasonably fast. Also, for polarization imaging, Clark searches for the peak in $I^2 + Q^2 + U^2 + V^2$.

Inside the Toolkit:

The `im.clean` method is used for CLEANing data. There are a number of methods used to set up the clean, including `im.setoptions`.

5.3.1.2 The hogbom algorithm

The hogbom algorithm is the “Classic” image-plane CLEAN, where model pixels are found iteratively by searching for the peak. Each point is subtracted from the full residual image using the shifted and scaled point spread function. In general, this is not a good choice for most imaging problems (`clark` or `csclean` are preferred) as it does not calculate the residuals accurately. But in some cases, with poor uv-coverage and/or a PSF with bad sidelobes, the Hogbom algorithm will do better as it uses a smaller beam patch. For polarization cleaning, Hogbom searches for clean peak in I , Q , U , and V independently.

5.3.1.3 The clarkstokes algorithm

In the `'clarkstokes'` algorithm, the Clark psf (§ 5.3.1.1) is used, but for polarization imaging the Stokes planes are cleaned sequentially for components instead of jointly as in `'clark'`. This means that this is the same as `'clark'` for Stokes I imaging only. This option can also be combined with `imagermode='csclean'` (§ 5.3.4).

5.3.2 The multiscale parameter

To activate multi-scale mode, specify a non-blank list of scales in the `multiscale` parameter. A good rule of thumb for starters is `[0, 2xbeam, 5xbeam]`, and maybe adding larger scales up to the maximum scale the interferometer can image. E.g. for a 2 arcsecond beam

```
multiscale = [0,6,10,30]    # Four scales including point sources
```

These are given in numbers of pixels, and specify FWHM of the Gaussians used to compute the filtered images.

Setting the `multiscale` parameter to a non-empty list opens up the sub-parameter:

```
multiscale      = [0, 6, 10, 30]    # set deconvolution scales (pixels)
  negcomponent  =          -1        # Stop cleaning if the
                                     # largest scale finds this number of neg
                                     # components
  smallscalebias =          0.6      # a bias to give more weight
                                     # toward smaller scales
```

The `negcomponent` sub-parameter is here to set the point at which the clean terminates because of negative components. For `negcomponent > 0`, component search will cease when this number of

Inside the Toolkit:

The `im.setscales` method sets the multi-scale Gaussian widths. In addition to choosing a list of sizes in pixels, you can just pick a number of scales and get a geometric series of sizes.

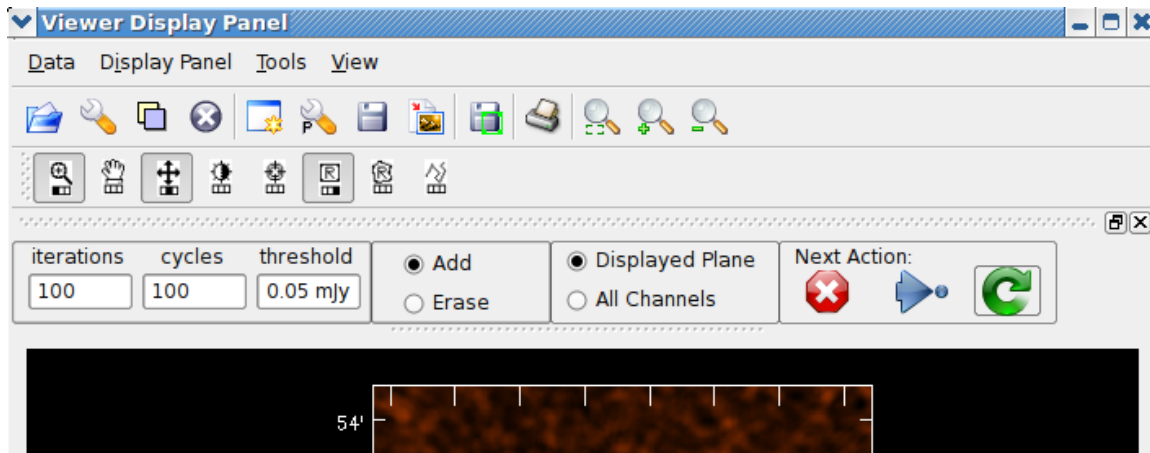


Figure 5.1: Close-up of the top of the interactive `clean` window. Note the boxes at the left (where the `iterations`, `cycles`, and `threshold` can be changed), the buttons that control add/erase, the application of mask to channels, and whether to stop, complete, or continue cleaning, and the row of Mouse-button tool assignment icons.

```

cyclefactor      =      1.5      # Controls how often major
                                # cycles are done. (e.g. 5 for
                                # frequently)
cyclespeedup     =      -1      # Cycle threshold doubles in
                                # this number of iterations

```

These options are explained below. In the CS mode, cleaning is split into minor and major cycles. For each field, a minor cycle is performed using the PSF algorithm specified in `psfmode` (§ 5.3.1). At major-cycle breakpoints, the points thus found are subtracted from the original visibilities. A fast variant does a convolution using a FFT. This will be faster for large numbers of visibilities. If you want to be extra careful, double the image size from that used for the Clark clean and set a mask to clean only the inner quarter or less (this is not done by default). This is probably the best choice for high-fidelity deconvolution of images without lots of large-scale structure.

Note that when using the Cotton-Schwab algorithm with a `threshold` (§ 5.3.12), there may be strange behavior when you hit the threshold with a major cycle. In particular, it may be above threshold again at the start of the next major cycle. This is particularly noticeable when cleaning a cube, where different channels will hit the threshold at different times.

In the empty mode (`imagermode=''`), the major and minor clean cycles work off of the gridded FFT dirty image, with residuals updated using the PSF calculation algorithm set by the `psfmode` parameter (§ 5.3.1). This method is not recommended for high dynamic range or high fidelity imaging applications, but can be significantly faster than CS clean (the default). Note that for this option only, if `mask=''` (no mask or box set) then it will clean the inner quarter of the image by default.

ALERT: You will see a warning message in the logger, similar to this:

```
Zero Pixels selected with a Flux limit of 0.000551377 and a maximum Residual of 0.00751239
```

whenever it find 0 pixels above the threshold. This is normal, and not a problem, if you've specified a non-zero threshold. On the other hand, if you get this warning with the threshold set to the default of '0Jy', then you should look carefully at your inputs or your data, since this usually means that the masking is bad.

The option `imagermode='mosaic'` is for multi-field mosaics. This choice opens up the sub-parameters

```
imagermode      = 'mosaic' # Use csclean or mosaic. If '', use psfmode
mosweight       = False   # Individually weight the fields of the mosaic
ftmachine       = 'mosaic' # Gridding method for the image
scaletype       = 'SAULT'  # Controls scaling of pixels in the image plane.
cyclefactor     = 1.5     # change depth in between of csclean cycle
cyclespeedup    = -1      # Cycle threshold doubles in this number of iteration
```

These options are explained below.

5.3.4.1 Sub-parameter `cyclefactor`

This sub-parameter is activated for `imagermode='csclean'` and `'mosaic'`.

The `cyclefactor` parameter allows the user to change the threshold at which the deconvolution cycle will stop and then degrid and subtract the model from the visibilities to form the residual. This is with respect to the breaks between minor and major cycles that the `clean` part would normally force. Larger values force a major cycle more often.

This parameter in effect controls the threshold used by CLEAN to test whether a major cycle break and reconciliation occurs:

$$\text{cycle threshold} = \text{cyclefactor} * \text{max sidelobe} * \text{max residual}$$

If `mosaic` or `csclean` diverges on your data, try a larger `cyclefactor`. A larger value typically increases the robustness of your deconvolution. The price, however, will be a slower algorithm. On the other hand, if you find that the cleaning is slow due to taking too many major cycle breaks, then reduce `cyclefactor`.

Note that currently the `cycle_threshold` will saturate at a maximum value of 0.80 even when you set `cyclefactor` to a very high value or you have very high PSF sidelobes. This means that with a `gain = 0.1` you will get 3 minor cycles per major cycle when hitting the limit.

Inside the Toolkit:

The `im.setmfcontrol` method sets the parameters that control the cycles and primary beam used in mosaicing.

Some rules of thumb:

If you have data taken with a small number of antennas, for example from ALMA in the commissioning and early-science phase, then you will have high sidelobes in the PSF. In this case, you will have to reduce `cyclofactor` considerably, likely into the range 0.25 to 0.5, if you want efficient cleaning of simple source structures (e.g. point sources). You can use the `viewer` to look at your PSF image and see what the maximum sidelobe level is and judge accordingly.

However, if your uv-coverage results in a poor PSF and you have complex source structure, then you should reconcile often (a `cyclofactor` of 4 or 5). For reasonable PSFs, use `cyclofactor` in the range 1.5 to 2.0. For good PSFs, or for faster cleaning at the expense of some fidelity, we recommend trying a lower value, e.g. `cyclofactor = 0.25`, which at least in some of our mosaicing tests led to a speedup of around a factor of two with little change in residuals.

5.3.4.2 Sub-parameter `cyclespeedup`

This sub-parameter is activated for `imagermode='csclean'` and `'mosaic'`.

The `cyclespeedup` parameter allows the user to let `clean` to raise the threshold at which a major cycle is forced if it is not converging to that threshold. To do this, set `cyclespeedup` to an integer number of iterations at which if the threshold is not reached, the threshold will be doubled. See `cyclofactor` above for more details. By default this is turned off (`cyclespeedup = -1`). In our tests, a value like `cyclespeedup = 50` has been used successfully.

5.3.4.3 Sub-parameter `ftmachine`

This sub-parameter is activated for `imagermode='mosaic'`. **Note:** The actual “ftmachine” used may be overridden by choices made to other parameters, such as `gridmode`.

The `ftmachine` parameter controls the gridding method and kernel to be used to make the image. A string value type is expected. Choices are: `'ft'`, `'sd'`, `'both'`, or `'mosaic'` (the default).

The `'ft'` option uses the standard gridding kernel (as used in `clean`).

The `'sd'` option forces gridding as in single-dish data.

For combining single-dish and interferometer MS in the imaging, the `'both'` option will allow `clean` to choose the `'ft'` or `'sd'` machines as appropriate for the data.

The `'mosaic'` option (the default) uses the Fourier transform of the frequency-dependent primary beam (the aperture cross-correlation function in the uv-plane) as the gridding kernel. This allows the data from the multiple fields to be gridded down to a single uv-plane, with a significant speed-up in performance in most (non-memory limited) cases. The effect of this extra convolution is an addi-

Inside the Toolkit:

The `im.setoptions` method sets the parameters relevant to mosaic imaging, such as the `ftmachine`.

tional multiplication (apodization) by the primary beam in the image plane. This can be corrected for, but does result in an image with optimal signal to noise ratio across it.

The primary beams used in CASA are described in § 5.2.14.

ALERT: Note that making a non-square image (e.g. using unequal values in `imsize`) for `ftmachine='mosaic'` will grid the data into a uv-plane with correspondingly non-square cells. This has not been extensively tested, and may result in undesired image artifacts. We recommend that the user make square mosaic images when possible, but in principle non-square images should work.

5.3.4.4 Sub-parameter `mosweight`

If `mosweight=False` (default) then the data will be weighted for optimal signal to noise ratio across the mosaic image. This should be used for most cases.

If `mosweight=True` then individual mosaic field relative weights will be readjusted on a per visibility basis (much like uniform gridding weights). This may give better performance in cases where one or a few fields in the mosaic have drastically different weights and/or integration time, and it is desired that the noise be more “uniform” across the mosaic image. Use this with care, we have not explored its use fully.

5.3.4.5 Sub-parameter `scaletype`

The `scaletype` parameter controls weighting of pixels in the image plane. This sub-parameter is activated for `imagermode='mosaic'`.

The default `scaletype='PBCOR'` scales the image to have the correct flux scale across it (out to the beam level cutoff `minpb`). This means that the noise level will vary across the image, being increased by the inverse of the weighted primary beam responses that are used to rescale the fluxes.

This option should be used with care, particularly if your data has very different exposure times (and hence intrinsic noise levels) between the mosaic fields.

If `scaletype='SAULT'` then the image will be scaled so as to have constant noise across it. This means that the point source response function varies across the image attenuated by the weighted primary beam(s). However, this response is output in the `.flux` image and can be later used to correct this.

Note that this scaling as a function of position in the image occurs after the weighting of mosaic fields specified by `mosweight` and implied by the gridding weights (`ftmachine` and `weighting`).

5.3.4.6 The threshold revisited

For mosaics, the specification of the threshold is not straightforward, as it is in the single field case. This is because the different fields can be observed to different depths, and get different weights in

Inside the Toolkit:

The `im.setmfcontrol` method gives more options for controlling the primary beam and noise across the image.

the mosaic. We now provide internal rescaling (based on `scaletype`) so `clean` does its component search on a properly weighted and scaled version of the sky.

For `ftmachine='ft'`, the minor cycles of the deconvolution are performed on an image that has been weighted to have constant noise, as in `'SAULT'` weighting (see § 5.3.4.5). This is equivalent to making a dirty mosaic by coadding dirty images made from the individual pointings with a sum of the mosaic contributions to a given pixel weighted by so as to give constant noise across the image. This means that the flux scale can vary across the mosaic depending on the effective noise (higher weighted regions have lower noise, and thus will have higher “fluxes” in the `'SAULT'` map). Effectively, the flux scale that threshold applies to is that at the center of the highest-weighted mosaic field, with higher-noise regions down-scaled accordingly. Compared to the true sky, this image has a factor of the PB, plus a scaling map (returned in the `.flux` image). You will preferentially find components in the low-noise regions near mosaic centers.

When `ftmachine='mosaic'` and `scaletype='SAULT'`, the deconvolution is also performed on a “constant noise image”, as detailed above for `'ft'`.

ALERT: The intrinsic image made using `ftmachine='mosaic'` is equivalent to a dirty mosaic that is formed by coadding dirty images made from the individual fields after apodizing each by the PB function. Thus compared to the true sky, this has a factor of the PB² in it. You would thus preferentially find components in the centers of the mosaic fields (even more so than in the `'ft'` mosaics). We now rescale this image internally at major-cycle (and interactive) boundaries based on `scaletype`, and do not have a way to clean on the raw unscaled dirty image (as was done in previous released versions).

5.3.5 Parameter interactive

If `interactive=True` is set, then an interactive window will appear at various “cycle” stages while you clean, so you can set and change mask regions. These breakpoints are controlled by the `npercycle` sub-parameter which sets the number of iterations of clean before stopping.

```
interactive    =   True    # use interactive clean (with GUI viewer)
npercycle     =    100    # Number of iterations before interactive prompt
```

Note that `npercycle` is currently the only way to control the breakpoints in interactive clean.

For spectral cube imaging, it is often easier to deal with each channel in turn, rather than cleaning all channels in each cycle. We therefore provide the `chaniter=True` option under `'mode'`, where it will clean a channel fully before moving to the next one. You will set masks for each channel.

See the example of interactive cleaning in § 5.3.14.

5.3.6 Parameter mask

The `mask` parameter takes a list of elements, each of which can be a list of coordinates specifying a box, or a string pointing to the name of a cleanbox file, mask image, or region file. These are used by CLEAN to define a region to search for components.

Note that for `imagermode=''` (§ 5.3.4) the default with `mask=''` is to restrict `clean` to the inner quarter of the image.

5.3.6.1 Setting clean boxes

`mask` can be a list of CASA regions. For example,

```
mask = 'box[[80pix, 80pix],[120pix,120pix]],circle[[150pix,150pix],10pix]'
```

defines a box and a circle. They will be applied to all channels. To define different regions for different channel ranges, it will be best to use interactive mode in `clean`, the `viewer` (note that the `viewer` still created old format regions - they are still supported in CASA 3.3) or to create a CASA region file that contain the different regions. Chapter D describes the syntax of CASA regions. They can be specified by;

```
mask = 'regionfile.rgn, regionfile2.rgn'
```

5.3.6.2 Using clean mask images

You can give the `mask` parameter a string containing the name of a mask image to be used for CLEAN to search for components. You can use `interactive=True` to create such a mask for your image (§ 5.3.5).

5.3.7 Parameter minpb

The `minpb` parameter sets the level down to which the primary beam (or more correctly the voltage patterns in the array) can go and have a given pixel included in the image. This is important as it defines where the edge of the visible image or mosaic is. The default is 0.1 or equivalent to the 10% response level. If there is a lot of emission near the edge, then set this lower if you want to be able to clean it out.

NOTE: The `minpb` parameter is the level in the “primary beam” (PB) at which the cut is made. If you are using `ftmachine='mosaic'` (§ 5.3.4.3), this will show up in the `.flux.pbcoverage` image (new in version 2.4.0). See the discussion of `threshold` (§ 5.3.4.6) for related issues.

5.3.8 Parameter modelimage

The `modelimage` parameter specifies the name(s) of one or more input starting image(s) to use to calculate the first residual before cleaning. These are used in addition to any image with a name defaulting from the `imagenam` root (e.g. on a restart). The output model will contain this model plus clean components found during deconvolution.

If the units of the image are `Jy/pixel`, then this is treated as a model image.

If the units of the image are `Jy/beam` or `Jy` per solid angle, then this is treated as a “single-dish” image and rescaled by the resolution (in the `'beam'` image header keyword). Inclusion of the SD image here is superior to feathering it in later. See § 5.6 for more information on feathering.

5.3.9 Parameter `niter`

The `niter` parameter sets the maximum total number of minor-cycle CLEAN iterations to be performed during this run of `clean`. If restarting from a previous state, it will carry on from where it was. Note that the `threshold` parameter can cause the CLEAN to be terminated before the requested number of iterations is reached.

5.3.10 Parameter `pbcor`

The `pbcor` parameter controls whether the final `.image` is scaled to correct for the Primary Beam of the array or not.

If `pbcor=False` (the default), then no such scaling is done and the image is in whatever “raw” scaling used by the `imagermode` algorithm underneath. For single-field cleaning with `imagermode=''` or `'csclean'`, this is the standard constant-noise image. If `imagermode='mosaic'`, then this is the `'SAULT'` scaled image (regardless of what `scaletype` is set to).

If `pbcor=True`, then at the end of deconvolution and imaging the “raw” image is rescaled by dividing by the noise and PB correction image. This is what is output by `clean` as the `.flux` image.

Note that regardless of what you set `pbcor` to, you can recover the other option using `immath` (§ 6.6) to either multiply or divide by the `.flux` image.

5.3.11 Parameter `restoringbeam`

The `restoringbeam` parameter allows the user to set a specific Gaussian restoring beam to make the final restored `.image` from the final `.model` and residuals.

If `restoringbeam=''` (the default), then the restoring beam is calculated by fitting to the PSF (e.g. the `.psf` image). For a mosaic, this is at the center of the field closest to the `phasecenter`.

The restoring beam can also be used to establish a single beam for large fractional bandwidths. If the PSF changes more than half a pixel across all channels in a cube, the PSF itself will be stored in the form of a cube, changing size from channel to channel. A specified restoring beam will output all planes at the same resolution and thus collapse to a single PSF (note that this can also be done in hindsight using `imsmooth`).

To specify a restoring beam, provide `restoringbeam` a list of [`bmaj`, `bmin`, `bpa`] which are the parameters of an elliptical Gaussian. The default units are in arc-seconds for `bmaj`, `bmin` components and degrees for the `bpa` component.

For example,

```
restoringbeam=['10arcsec']           # circular Gaussian FWHM 10"
restoringbeam=['10.0','5.0','45.0deg'] # 10"x5" at PA=45 degrees
```

5.3.12 Parameter threshold

The `threshold` parameter instructs `clean` to terminate when the maximum absolute residual reaches this level or below. Note that it may not reach this residual level due to the value of the `niter` parameter which may cause it to terminate early.

If `threshold` is given a floating-point number, then this is the threshold in milli-Jansky.

You can also supply a flux density *quanta* to `threshold`, e.g.

```
threshold = '8.0mJy'
threshold = '0.008Jy'
```

(these do the same thing).

5.3.13 Parameter gridmode

The `gridmode` parameter is now provided to access more advanced deconvolution capabilities. The default `gridmode=''` is recommended for most cases.

The `gridmode='widefield'` option allows imaging in the wide-field regime where the W -term is not negligible. The CASA implementation allows both standard uv -plane faceting as well as the W -Projection algorithm² or a combination of the two. Its sub-parameters are:

```
gridmode      = 'widefield' # Gridding kernel for FFT-based
                # transforms, default='' None
wprojplanes   = 1          # Number of w-projection planes for convolution
facets        = 1          # Number of facets along each axis (main image only)
```

The `wprojplanes` parameter sets the number of pre-computed w -planes used for the W -Projection algorithm (`wprojplanes=1` disables w -projection). The `facets` parameter sets the number of facets used. W -Projection, if used, is done for each facet. See § 5.3.18 below for more on wide-field imaging.

`gridmode='aprojection'`: A-Projection is an algorithm to account for the effects of the antenna primary beam (PB) during imaging. The time-dependent effects of the PB are projected-out during the imaging phase and the PB is included in the prediction phase of the iterative image deconvolution (see Bhatnagar, Cornwell, Golap & Uson 2008, A&A, 487, 419) for more details. Please also refer to this publication in your papers if this algorithm is used for imaging. The narrow-band A-Projection can be used by setting the `gridmode='aprojection'` in the `clean` task. This opens up the following new parameters:

²Cornwell et al. IEEE JSTSP (2008).

```

gridmode   = 'aprojection'   # Gridding kernel for FFT-based
                                # transforms, default='' None
cfcache    = 'cfcache.dir'   # Convolution function cache directory
rotpainc   =      5.0        # Parallaxtic angle increment
                                # (degrees) for OTF A-term rotation
painc      =      360.0      # Parallaxtic angle increment
                                # (degrees) for computing A-term

```

`cfcache` is used to cache functions required in the A-Projection algorithm. The PB is rotated on-the-fly when a change of greater than `rotpainc` is detected. Alternatively, PB is re-computed if the P.A. changes by greater than `painc`.

Note that this code is still in the development and testing stage and should be used on shared-risk basis. Note also that the cost of imaging will be higher when using A-Projection. Therefore make a careful evaluation of whether you need to invoke it.

5.3.14 Interactive Cleaning — Example

If `interactive=True` is set, then an interactive window will appear at various “cycle” stages while you clean, so you can set and change mask regions. These breakpoints are controlled by the `npercyc` sub-parameter which sets the number of iterations of clean before stopping.

The window controls are fairly self-explanatory. It is basically a form of the `viewer`. A close-up of the controls are shown in Figure 5.1, and an example can be found in Figures 5.2–5.4. You assign one of the drawing functions (rectangle or polygon, default is rectangle) to the right-mouse button (usually), then use it to mark out regions on the image. Zoom in if necessary (standard with the left-mouse button assignment). Double-click inside the marked region to add it to the mask. If you want to reduce the mask, click the **Erase** radio button (rather than **Add**), then mark and select as normal. When finished setting or changing your mask, click the green clockwise arrow “Continue Cleaning” Next Action button. If you want to finish your clean with no more changes to the mask, hit the blue right arrow “Apply mask edits and proceed with non-interactive clean” button. If you want to terminate the clean, click the red X “Stop deconvolving now” button.

While stopped in an interactive step, you can change a number of control parameters in the boxes provided at the left of the menu bar. The main use of this is to control how many `iterations` before the next breakpoint (initially set to `npercyc`), how many cycles before completion (initially equal to `niter/npercyc`), and to change the `threshold` for ending cleaning. Typically, the user would start with a relatively small number of iterations (50 or 100) to clean the bright emission in tight mask regions, and then increase this as you get deeper and the masking covers more of the emission region. For extended sources, you may end up needing to clean a large number of components (10000 or more) and thus it is useful to set `niter` to a large number to begin with — you can always terminate the clean interactively when you think it is done. Note that if you change `iterations` you may also want to change `cycles` or your clean may terminate before you expect it to.

For strangely shaped emission regions, you may find using the polygon region marking tool (the second from the right in the button assignment toolbar) the most useful.

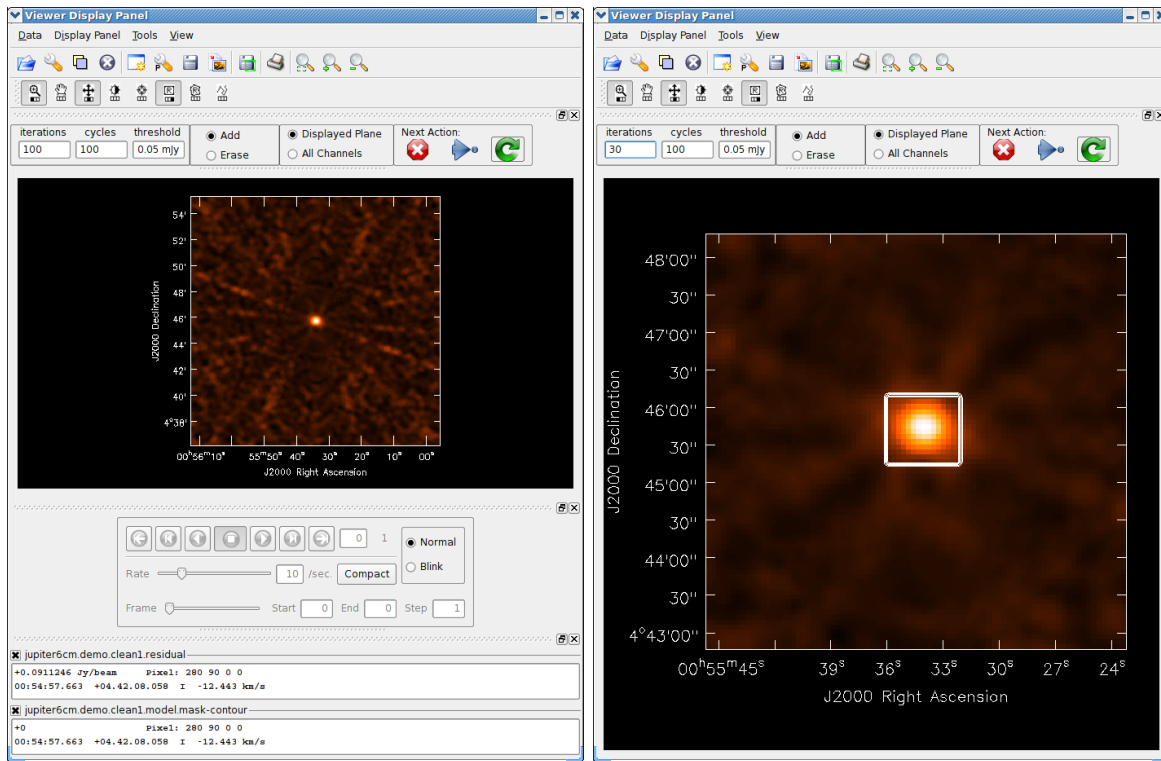


Figure 5.2: Screen-shots of the interactive `clean` window during deconvolution of the VLA 6m Jupiter dataset. We start from the calibrated data, but before any self-calibration. In the initial stage (left), the window pops up and you can see it dominated by a bright source in the center. Next (right), we zoom in and draw a box around this emission. We have also at this stage dismissed the tape deck and Position Tracking parts of the display (§ 7.2) as they are not used here. We have also changed the `iterations` to 30 for this boxed clean. We will now hit the Next Action **Continue Cleaning** button (the green clockwise arrow) to start cleaning.

The sequence of cleaning starting with the “raw” externally calibrated data is shown in Figures 5.2 – 5.4.

The final result of all this cleaning for Jupiter is shown in Figure 5.5. The `viewer` (§ 7) was used to overplot the polarized intensity contours and linear polarization vectors calculated using `immath` (§ 6.6) on the total intensity. See the following chapters on how to make the most of your imaging results.

For spectral cube images you can use the tape deck to move through the channels. You also use the panel with radio buttons for choosing whether the mask you draw applies to the `Displayed Plane` or to `All Channels`. See Figure 5.6 for an example. Note that currently the `Displayed Plane` option is set by default. This toggle is unimportant for single-channel images or `mode='mfs'`.

Advanced Tip: Note that while in interactive `clean`, you are using the `viewer`. Thus, you have the ability to open and register other images in order to help you set up the clean mask. For

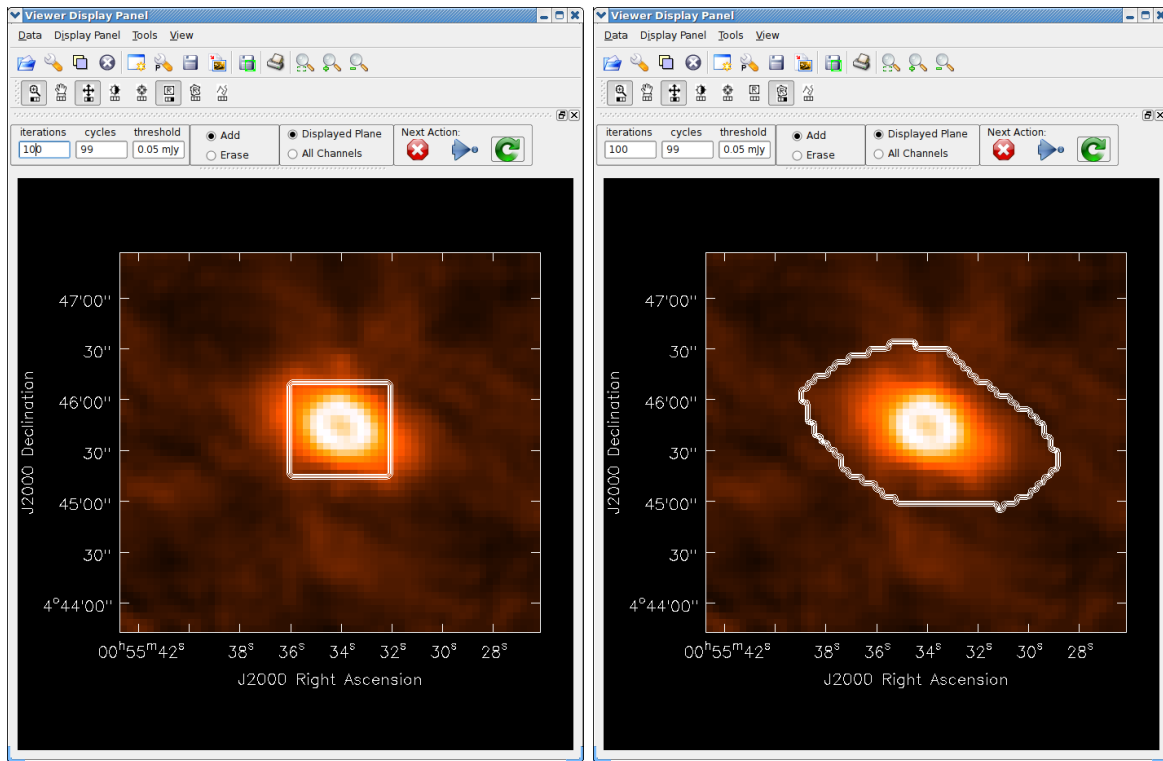


Figure 5.3: We continue in our interactive cleaning of Jupiter from where Figure 5.2 left off. In the first (left) panel, we have cleaned 30 iterations in the region previously marked, and are zoomed in again ready to extend the mask to pick up the newly revealed emission. Next (right), we have used the Polygon tool to redraw the mask around the emission, and are ready to **Continue Cleaning** for another 100 iterations.

example, if you have a previously cleaned image of a complex source or mosaic that you wish to use to guide the placement of boxes or polygons, just use the **Open** button or menu item to bring in that image, which will be visible and registered on top of your dirty residual image that you are cleaning on. You can then draw masks as usual, which will be stored in the mask layer as before. Note you can blink between the new and dirty image, change the colormap and/or contrast, and carry out other standard viewer operations. See § 7 for more on the use of the viewer.

ALERT: Currently, interactive spectral line cleaning is done globally over the cube, with halts for interaction after searching all channels for the requested `npercycle` total iterations. It is more convenient for the user to treat the channels in order, cleaning each in turn before moving on. This will be implemented in an upcoming update.

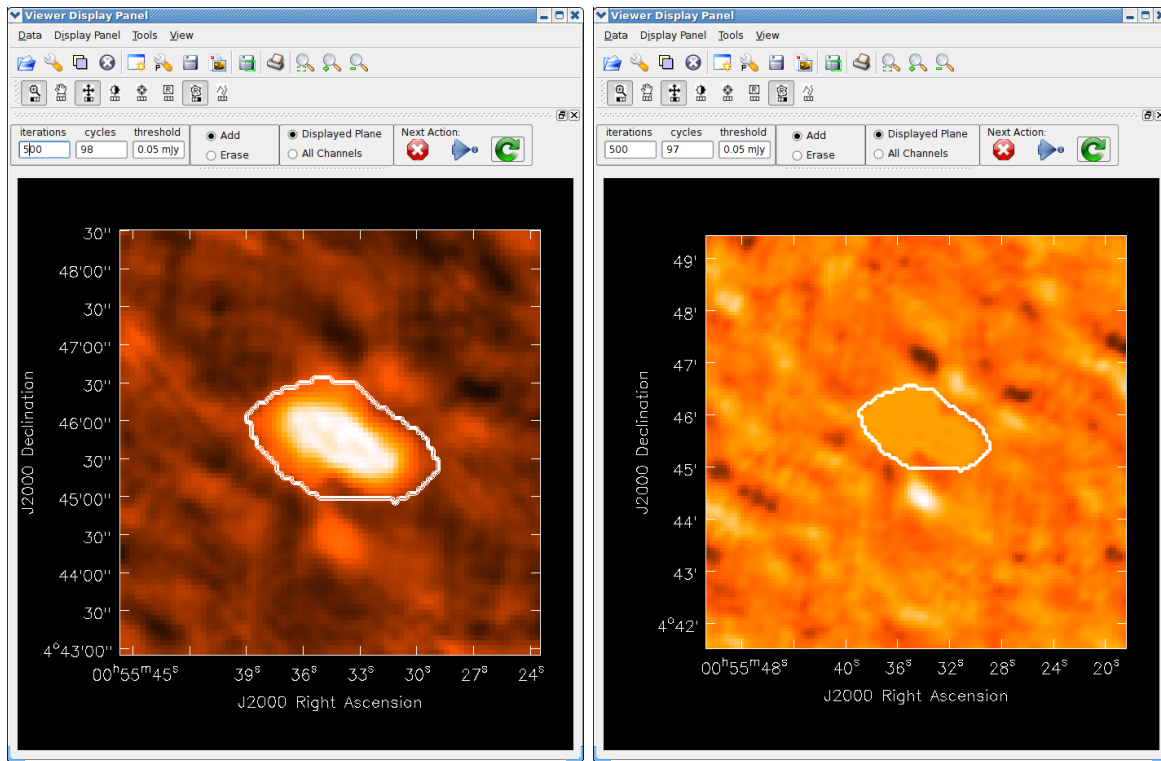


Figure 5.4: We continue in our interactive cleaning of Jupiter from where Figure 5.3 left off. In the first (left) panel, it has cleaned deeper, and we come back and zoom in to see that our current mask is good and we should clean further. We change `npercycle` to 500 (from 100) in the box at upper right of the window. In the final panel (right), we see the results after this clean. The residuals are such that we should terminate the `clean` using the red X button and use our model for self-calibration.

5.3.15 Mosaic imaging

The `clean` task contains the capability to image multiple pointing centers together into a single “mosaic” image. This ability is controlled by setting `imagermode='mosaic'` (§ 5.3.4).

The key parameter that controls how `clean` produces the mosaic is the `ftmachine` sub-parameter (§ 5.3.4.3). For `ftmachine='ft'`, `clean` will perform a weighted combination of the images produced by transforming each mosaic pointing separately. This can be slow, as the individual sub-images must be recombined in the image plane. **NOTE:** this option is preferred for data taken with sub-optimal mosaic sampling (e.g. fields too far apart, on a sparse irregular pattern, etc.).

The primary beams used in CASA are described in § 5.2.14.

If `ftmachine='mosaic'`, then the data are gridded onto a single uv-plane which is then transformed to produce the single output image. This is accomplished by using a gridding kernel that approximates the transform of the primary beam pattern. Note that for this mode the

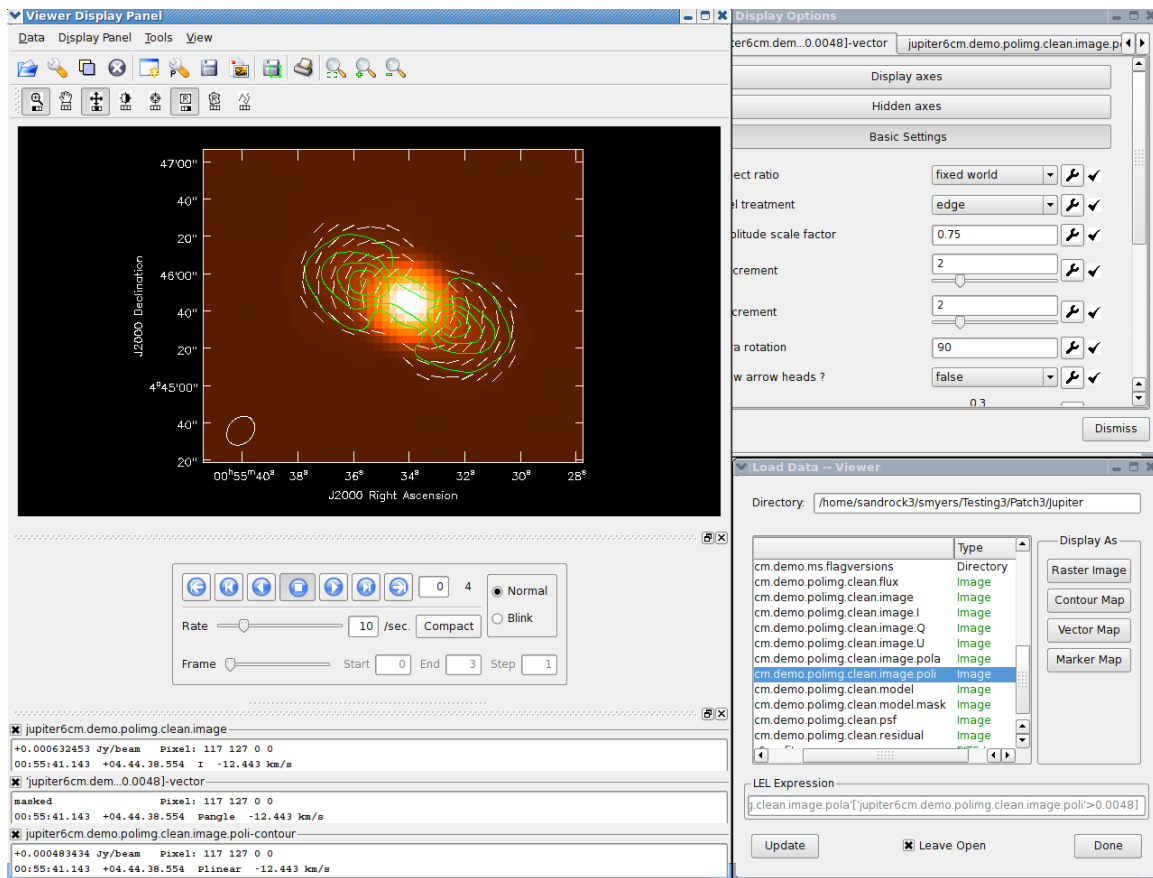


Figure 5.5: After clean and self-calibration using the intensity image, we arrive at the final polarization image of Jupiter. Shown in the viewer superimposed on the intensity raster is the linear polarization intensity (green contours) and linear polarization B-vectors (vectors). The color of the contours and the sampling and rotation by 90 degrees of the vectors was set in the Display Options panel. A LEL expression was used in the Load Data panel to mask the vectors on the polarized intensity.

`<imagename>.flux` image includes this convolution kernel in its effective weighted response pattern (needed to “primary-beam correct” the output image). For this mode only, an additional image `<imagename>.flux.pbcoverage` is produced that is the primary-beam coverage only used to compute the `minpb` cutoff (§ 5.3.7).

The `flatnoise` parameter determines whether the minor cycle performs on the the residual with or without a primary beam correction. Whereas the former has the correct fluxes, the latter has a uniform noise, which allows for a simpler deconvolution in particular at the the edges of the mosaic where the primary beam correction is largest.

ALERT: In order to avoid aliasing artifacts for `ftmachine='mosaic'` in the mosaic image, due to the discrete sampling of the mosaic pattern on the sky, you should make an image in which the

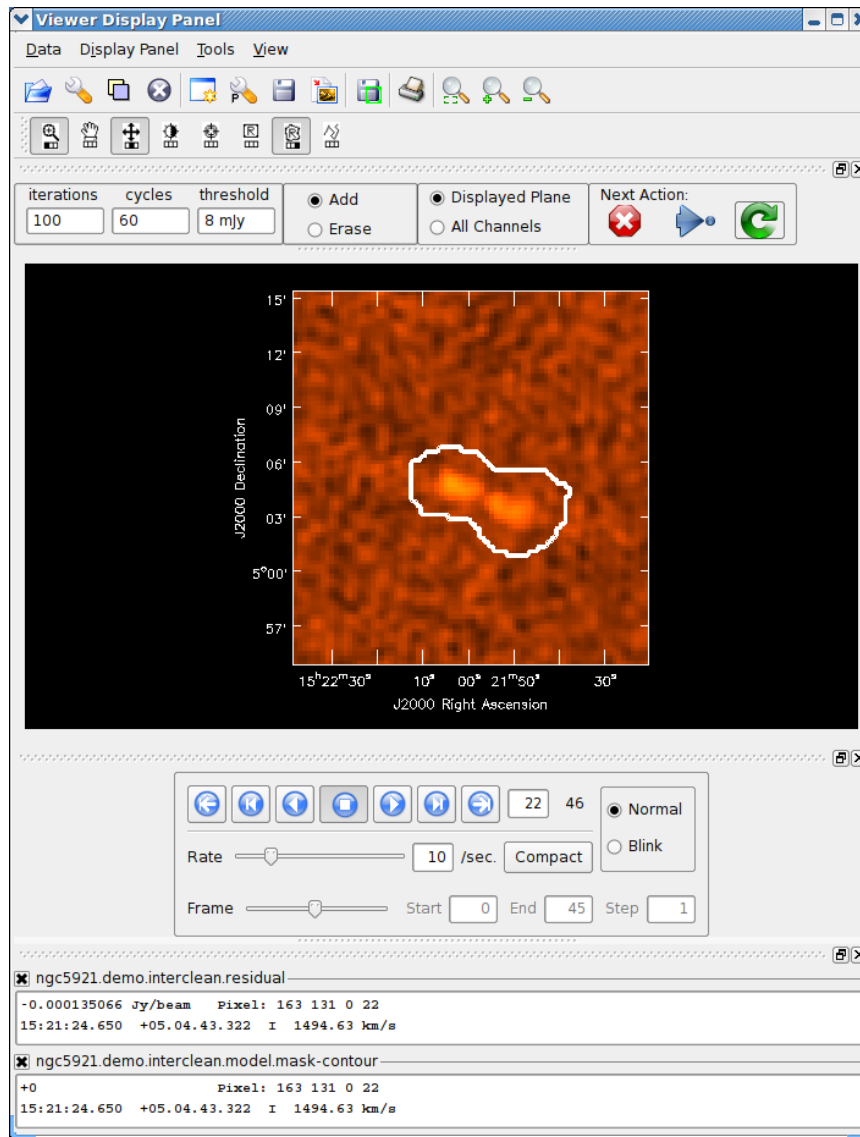


Figure 5.6: Screen-shot of the interactive `clean` window during deconvolution of the NGC5921 spectral line dataset. Note where we have selected the mask to apply to the `Displayed Plane` rather than `All Channels`. We have just used the Polygon tool to draw a mask region around the emission in this channel, which will apply to this channel only.

desired unmasked part of the image (above `minpb`) lies within the inner quarter. In other words, make an image twice as big as necessary to encompass the mosaic.

It is also important to choose an appropriate `phasecenter` for your output mosaic image (§ section:im.pars.phasecenter). The phase center should not be at the edge of an image with pointings around it. In that case, FFT aliasing may creep into the image.

An example of a simple mosaic `clean` call is shown below:

```
clean(vis='n4826_tboth.ms',
      imagename='tmosaic',
      mode='channel',
      nchan=30,start=46,           # Make the output cube 30 chan
      width=4,                   # start with 46 of spw 0, avg by 4 chans
      spw='0~2',
      field='0~6',
      cell=[1.,1.],
      imsize=[256,256],
      stokes='I',
      psfmode='clark',
      niter=500,
      imagermode='mosaic',
      scaletype='SAULT',
      cyclefactor=0.1)
```

5.3.16 Heterogeneous imaging

The `clean` task and underlying tools can handle cases where there are multiple dish sizes, and thus voltage patterns and primary beams, in the array. This is effected by using the dish sizes stored in the `ANTENNA` sub-table of the MS. Depending on how the data was written and imported into CASA, the user may have to manually edit this table to insert the correct dish sizes (e.g. using `browsetable` or the `tb` table tool).

5.3.17 Polarization imaging

The `clean` task handles full and partial Stokes polarization imaging through the setting of the `stokes` parameter (§ 5.2.10). The subsequent deconvolution of the polarization planes of the image and the search for clean components is controlled by the `psfmode` parameter (§ 5.3.1). If the `stokes` parameter includes polarization planes other than I, then choosing `psfmode='hogbom'` (§ 5.3.1.2) or `psfmode='clarkstokes'` (§ 5.3.1.3) will clean (search for components) each plane sequentially, while `psfmode='clark'` (§ 5.3.1.1) will deconvolve jointly.

The interactive `clean` example given above (§ 5.3.14) shows a case of polarization imaging.

5.3.18 Wide-field imaging and deconvolution in `clean`

When imaging sufficiently large angular regions, the sky can no longer be treated as a two-dimensional plane and the use of the standard `clean` task will produce distortions around sources that become increasingly severe with increasing distance from the phase center. In this case, one must use a “wide-field” imaging algorithm such as w-projection or faceting.

When is wide-field imaging needed? The number of required facets N depends on the maximum baseline B_{\max} , the dish diameter D and the wavelength λ as:

Receiver Band	Wavelength [cm]	Array configurations
4	430	A/B/C/D
L	20	A/B/C
S	10	A/B
C	5	A
X	3	A
Ku	2	A
K	1.4	–
Ka	0.9	–
Q	0.7	–

Table 5.1: Combinations of observing band (wavelength,) and antenna array configurations that require w-projection.

$$N = \frac{B_{\max} \lambda}{D^2} \quad (5.13)$$

and w-projection is required when $N > 1$. (For details, see “Synthesis Imaging in Radio Astronomy II”, ed. Taylor, G., Carilli, C., Perley, R. 1999). With 25 m diameter JVLA dishes (which implies that imaging is requested out to the primary beam FWHM), w-projection is required for array configurations as listed in Table 5.1.

The relevant inputs for `clean` for wide-field imaging are:

```

gridmode          = 'widefield'      # The kind gridding kernel to
                                     # be used for FFT-based transforms
  wprojplanes     =          1        # Number of w-projection planes for convolution
  facets          =          1        # Number of facets along each axis (main image only)

```

Most of the `clean` parameters behave as described previously.

Wide-field imaging can be carried out using two major modes: First, the w-projection mode as chosen with `ftmachine` deals with the w-term (the phase associated with the sky/array curvature) internally. Secondly, the image can be broken into many facets, each small enough so that the w-term is not significant. These two basic methods can be combined, as discussed below in § 5.3.18.4.

5.3.18.1 Outlier fields

When using wide-field imaging, the position and image size of any independent images must be specified. Those positions will be used to add additional cleaning components to strong sources that may reside in that area and influence the central image.

There are a two options to specify the outlier fields:

Direct listing of fields The outlier field directions are provided via their centers (`phasecenter` parameter), and their sizes as a second entry in the `imsize` parameter, e.g. 128 pixels in the

example below. `clean` will derive two additional images and their names are to be provided in the `imagename` field that will then be a list of the main field name plus the outlier field names:

```
vis           = 'wfield.ms'           # name of input visibility file
imagename=[ 'n5921', 'outlier1', 'outlier2' ] # Pre-name of output images
outlierfile  = ''                     # Text file with image names, sizes, centers
mask         = [['image_setup.rgn'], [''], ['']]
imsize       = [[2048,2048],[128,128],[128,128]] # Image size in pixels (nx,ny)
cell         = '1.0arcsec'           # The image cell size in arcseconds [x,y].
phasecenter  = ['', 'J2000 13h27m20.98 43d26m28.0', 'J2000 13h30m52.158 43d23m08.00']
```

Outlier file For many outlier fields, it may be easier to setup the main interface to `clean` for the main field only and list outlier fields in an additional `outlierfile`:

```
imagename='n5921'
outlierfile = 'outliers.txt'
imsize=[1024,1024]
phasecenter = ''
```

`outliers.txt` provides all outlier fields with a syntax that is similar to the direct input, but separated by field. Below is an example for an `outlierfile`:

```
#content of outliers.txt
#
#outlier field1
imagename='outlier1'
imsize=[512,512]
phasecenter = 'J2000 13h30m52.15 43d23m08.0'
mask='box[[245pix,245pix],[265pix,265pix]]'
#
#outlier field2
imagename='outlier2'
imsize=[512,512]
phasecenter = 'J2000 13h24m08.16 43d09m48.0'
```

The syntax rules for the outlier files are:

- each field must begin with `imagename` followed by
- `imsize` and `phasecenter` must be given
- optionally a `mask` can be provided. The `mask` parameter follows the CASA region file convention (Chapter D) or can be a mask file or LEL string.

The older AIPS-style convention (and box definition) that was used in CASA 3.2 and earlier is still supported in CASA 3.3 but will be deprecated for CASA 3.4 and higher.

5.3.18.2 Setting up w-projection

The w-projection mode is controlled using `wprojplanes` sub-parameter, e.g.

```
gridmode          = 'widefield'      # The kind gridding kernel to
                                # be used for FFT-based transforms
    wprojplanes    =          64      # Number of w-projection planes for convolution
    facets         =           1      # Number of facets along each axis (main image only)
```

will construct 64 w-projection planes.

The w-projection algorithm is much faster than using faceting, but it does consume a lot of memory. On most 32-bit machines with 1 or 2 Mbytes of memory, images larger than about 4000×4000 cannot be made.

5.3.18.3 Setting up faceting

Faceting will break the image into many small parts. This is invoked using `facets`:

```
gridmode          = 'widefield'      # The kind gridding kernel to be used for FFT-based transforms
    wprojplanes    =           1      # Number of w-projection planes for convolution
    facets         =           7      # Number of facets along each axis (main image only)
```

In this example the image is broken into 49 (7×7) facets.

A reasonable value of facets is such that the image width of each facet does not need the w-term correction. The computation method with pure faceting is slow, so that w-projection is recommended

5.3.18.4 Combination of w-projection and faceting

You can also use a combination of w-projection and faceting:

```
gridmode          = 'widefield'      # The kind gridding kernel to be used for FFT-based transforms
    wprojplanes    =          32      # Number of w-projection planes for convolution
    facets         =           3      # Number of facets along each axis (main image only)
```

This hybrid method allows for a smaller number of `wprojplanes` in order to try to conserve memory if the image size approached the memory limit of the computer. However, there is a large penalty in execution time.

5.4 Refactored clean: tclean

`tclean` is a refactored version of `clean` with a better interface, more possible combinations of algorithms and new algorithms. `tclean` also allows to process the imaging and deconvolution on parallelized computing methods. `tclean` is still undergoing testing and improvement and we refer to the inline help for a detailed task description.

Eventually `tclean` will replace the current `clean` task.

5.5 Primary Beam Correction (`impbcor`, `widebandpbcor`)

The primary beam correction can be applied during the imaging with `clean`. It is also possible to correct after imaging using the task `impbcor` for 'regular' data sets, or `widebandpbcorr` for those that used the Taylor-term expansion function in `clean` (`nterms>1`). `pbcor` has the following inputs:

```
# impbcor :: Construct a primary beam corrected image from an image
and a primary beam pattern.
image_name      =      ''      # Name of the input image
pbimage         =      ''      # Name of the primary beam
                                # image which must exist or
                                # array of values for the pb response. Default ""
outfile        =      ''      # Output image name. If empty, no image is written.
                                # Default ""
box            =      ''      # One or more boxes to use
                                # for fit region(s). Default is
                                # to use the entire directional plane.
    region     =      ''      # The region to correct. Default is entire image. If
                                # both box and region are specified, box is used and
                                # region is not.

chans          =      ''      # The frequency planes to correct. Default is all
                                # frequencies.
stokes         =      'I'     # The correlations to correct. Default is all.
mask          =      []      # Boolean LEL expression or mask region. Default is
                                # none.
mode          =      'velocity' # Divide or multiply the image by the primary beam
                                # image. Minimal match supported. Default "divide"
cutoff        =      -1.0     # PB cutoff. If mode is "d", all values less than this
                                # will be masked. If "m", all values greater will be
                                # masked. Less than 0, no cutoff. Default no cutoff
wantreturn    =      False    # Return an image tool
                                # referencing the corrected image?
```

The main inputs are the input image and the image of a primary beam (usually your “`image.flux`” output image from `clean`) in the `pbimage` parameter. The `mode` parameter will typically be `'divide'` but it is also possible to multiply with the beam pattern.

`widebandpbcor` has the following options

```

# widebandpbcor :: Wideband PB-correction on the output of the MS-MFS algorithm
vis          =      ''          # Name of measurement set.
imagename    =      ''          # Name-prefix of multi-termimages to operate on.
nterms       =      2           # Number of taylor terms to use
threshold    =      ''          # Intensity above which to
                                # re-calculate spectral index
action       =      'pbcor'     # PB-correction (pbcor) or
                                # only calc spectral-index (calcalpha)
    reffreq   =      ''          # Reference frequency (if specified in clean)
    pbmin     =      0.2         # PB threshold below which to not correct
    field     =      ''          # Fields to include in the PB calculation
    spwlist   =      ''          # List of N spw ids
    chanlist  =      []          # List of N channel ids
    weightlist =      []         # List of N weights (relative)

```

`action='pbcor'` computes Taylor-coefficient images that represent the primary beam spectrum and applies them to the input Taylor coefficient images. The `action='calcalpha'` will recalculate spectral index maps based on the primary beam correction polynomials.

5.6 Combined Single Dish and Interferometric Imaging (feather)

The term “feathering” is used in radio imaging to describe how to combine or “feather” two images together by forming a weighted sum of their Fourier transforms in the (gridded) uv -plane. Intermediate size scales are down-weighted to give interferometer resolution while preserving single-dish total flux density.

The feathering technique does the following:

1. The single-dish and interferometer images are Fourier transformed.
2. The beam from the single-dish image is Fourier transformed ($FTSDB(u, v)$), (alternatively, one can specify some smaller portion of the single dish aperture, which corresponds to a wider beam).
3. The Fourier transform of the interferometer image is multiplied by $(1 - FTSDB(u, v))$. This basically down weights the shorter spacing data from the interferometer image.
4. The Fourier transform of the single-dish image is scaled by the volume ratio of the interferometer restoring beam to the single dish beam.
5. The results from 3 and 4 are added and Fourier transformed back to the image plane.

The term feathering derives from the tapering or down-weighting of the data in this technique; the overlapping, shorter spacing data from the deconvolved interferometer

Other Packages:

The `feather` task is analogous to the AIPS `IMERG` task and the MIRIAD `immerge` task with option `'feather'`.

image is weighted down compared to the single dish image while the overlapping, longer spacing data from the single-dish are weighted down compared to the interferometer image.

The tapering uses the transform of the low resolution point spread function. This can be specified as an input image or the appropriate telescope beam for the single-dish. The point spread function for a single dish image may also be calculated using `clean`.

Advice: Note that if you are feathering large images, be advised to have the number of pixels along the X and Y axes to be composite numbers and definitely not prime numbers. In general FFTs work much faster on even and composite numbers. You may use `subimage` function of the `image` tool to trim the number of pixels to something desirable.

The inputs for `feather` are:

```
# feather :: Combine two images using their Fourier transforms
imagename      =      ''      # Name of output feathered image
highres        =      ''      # Name of high resolution (interferometer) image
lowres         =      ''      # Name of low resolution (single dish) image
sdfactor       =      1.0     # Scale factor to apply to Single Dish image
effdishdiam    =      -1.0    # New effective SingleDish diameter to use in m
lowpassfiltersd =      False  # Filter out the high spatial frequencies of the SD
                                     # image
```

The single-dish data cube is specified by the `lowres` and the interferometric data cube by the `highres` keyword. The combined, feathered output cube name is given by the `imagename` parameter. `sdfactor` can be used to adjust the flux calibration of the images. Since single-dish processing typically involves the fit of a baseline level, it might be the one with the most uncertain calibration and `sdfactor` will multiply with the single-dish image values for any needed correction.

The weighting functions for the data are usually the Fourier transform of the Single Dish beam $\text{FFT}(\text{PB}_{SD})$ for the Single dish data, and the inverse, $1\text{-FFT}(\text{PB}_{SD})$ for the interferometric data. It is possible, however, to change the weighting functions by pretending that the SD is smaller in size via the `effdishdiameter` parameter. This tapers the high spatial frequencies of the SD data and adds more weight to the interferometric data. The `lowpassfiltersd` can take out artifacts at very high spatial frequencies that are often present but non-physical in SD data.

Alert: Before feathering data, make sure that the single dish data header is correct. The units need to be in $Jy/beam$, and the beam to be in the usual $bmin$ and $bmaj$ keywords (check with `imhead` if CASA recognizes all header information correctly). `feather` also does not accept interferometric data cubes with a varying beam size per channel. Prior to the combination, make sure that the cube has a single beam size for all channels, you may use `imsmooth` with `kernel='commonbeam'` to smooth the data to the largest beam in the cube.

Note that the only inputs are for images and `feather` will attempt to regrid the images to a common shape, i.e. pixel size, pixel numbers, and spectral channels. `feather` does not do any

deconvolution but combines presumably deconvolved images after the fact. This implies that the short spacings extrapolated by the deconvolution process will be those that are down-weighted the most. The single dish image must have a well-defined beam shape and the correct flux units for a model image (Jy/beam instead of Jy/pixel) so use the tasks `imhead` and `immath` first to convert if needed.

Starting with a cleaned synthesis image and a low resolution image from a single dish telescope, the following example shows how they can be feathered:

```
feather(imagename='feather.im',      # Create an image called feather.im
        highres='synth.im',         # The synthesis image is called synth.im
        lowres='single_dish.im'     # The SD image is called single_dish.im
    )
```

5.6.1 Visual Interface for feather (casafeather)

CASA also provides a visual interface to the `feather` task. The interface is run *from a command line outside CASA* by typing `casafeather` in a shell. Fig. 5.7 shows an example. As a first step, one needs to specify a high and a low resolution image, typically an interferometric and a single dish map. Note that the single dish map needs to be in units of Jy beam⁻¹. An output image is usually specified, too, and an additional image, such as a non-deconvolved (dirty) interferometric image can be specified, too. On the main GUI, press “Feather” to start the feathering process, which includes regridding the low resolution image to the high resolution image.

“casafeather” has the ability to show two major rows of displays (see Fig. 5.7) that can be turned on or off. A good visualization is usually obtained by making both axes logarithmic. This can be specified in the “Customize menu”, the *toothed wheel* symbol at the top of the panel. The two rows of displays are: 1) “Original Data Slice”: Cuts through the *u* and *v* directions of the Fourier transformed input images. A vertical line shows the location of the effective dish diameter(s). 2) “Feathered Data Slice”: The same cuts, but scaled by the “low resolution scale factor” and weighted by the weighting functions (see § 5.6). In this display, the weighting functions themselves are shown, too.

At the top of the display `effdshdiameter` for *u* and *v* and `sdfactor` can be provided in the “Effective Dish Diameter” and “Low Resolution Scale Factor” input boxes.

The data can be visualized in different forms. The data type to be displayed can be selected in the “Color Options” menu. The data can be the unmodified, original data, or data that have been convolved with the high or low resolution beams. One can also select to display data that were weighted and scaled by the functions discussed above.

The data can also be displayed in the form of a “scatter plot” (Fig. 5.8). This allows one to check for differences in flux. In particular, the scaling parameter should be adjusted such that the flux of the Low-resolution data, convolved with the High beam, weighted and scaled, is the same as the Dirty data, convolved with the Low beam, weighted (use the High resolution data instead of the Dirty data if the latter are not available). If that can be achieved, the flux adjustments should be roughly correct. The “scatter plot” can display any two data sets on the two axes, selected from the “Color Preferences” menu.

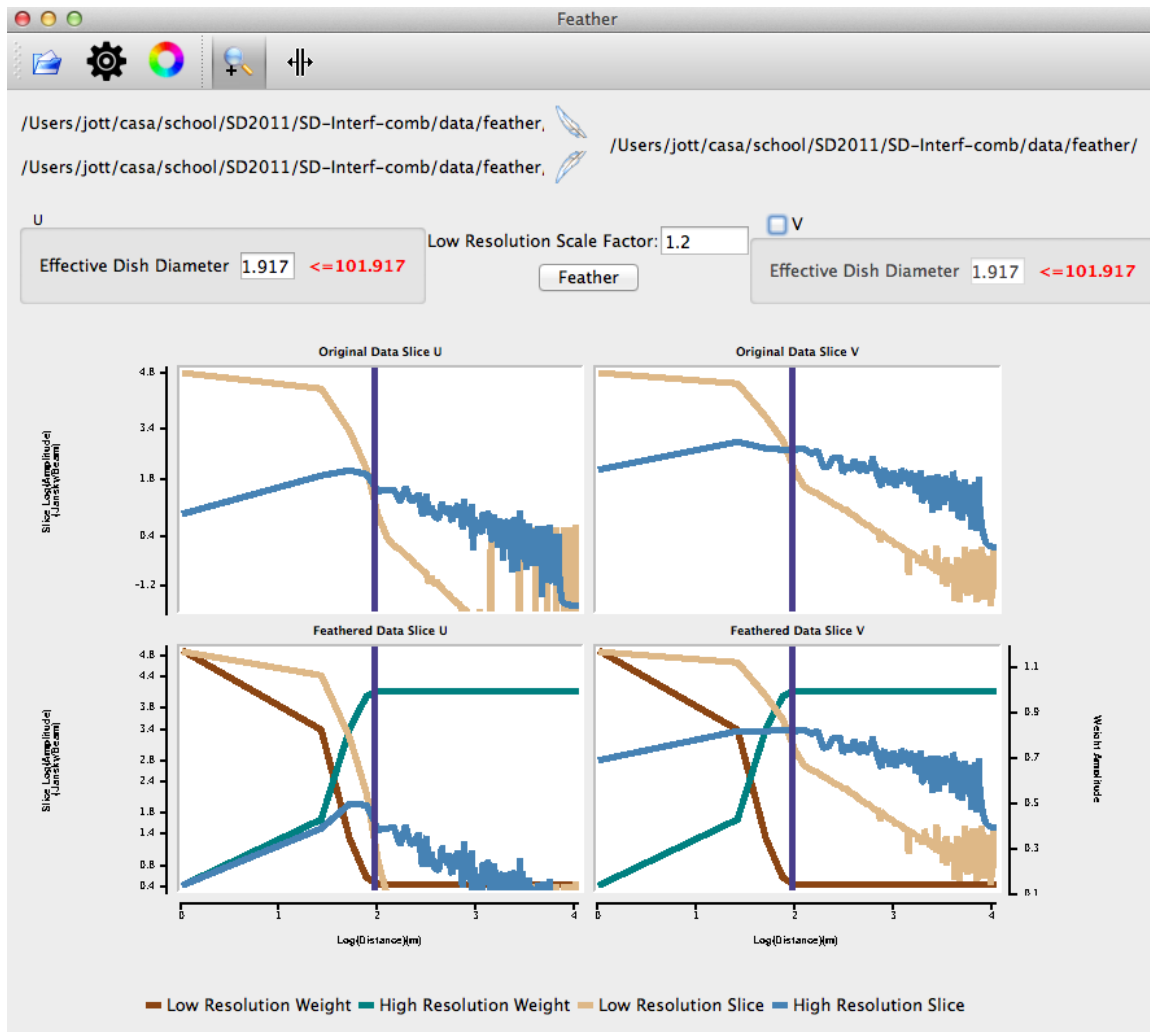


Figure 5.7: Visual “casafeather” interface to the feather task.

The “Customize” button at the top (toothed wheel), allows one to set the display parameters as seen in Fig. 5.9. Options are to show the slice plot, the scatter plot, or the legend. One can also select between logarithmic and linear axes, and whether the x-axis for the slices are in the u, or v, or both directions, or, alternatively a radial average in the uv-plane can be used. For data cubes one can also select a particular velocity plane, or to average the data across all velocity channels.

5.7 Making Deconvolution Masks or Box Regions

For most careful imaging, you will want to restrict the region over which you allow CLEAN components to be found. To do this, you can create a ‘deconvolution region’ or ‘mask’ image using the

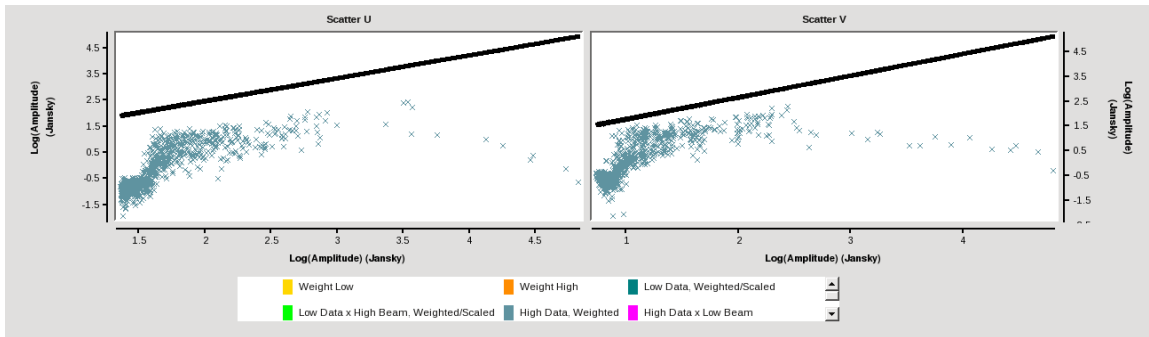


Figure 5.8: The scatter plot in casafeather.

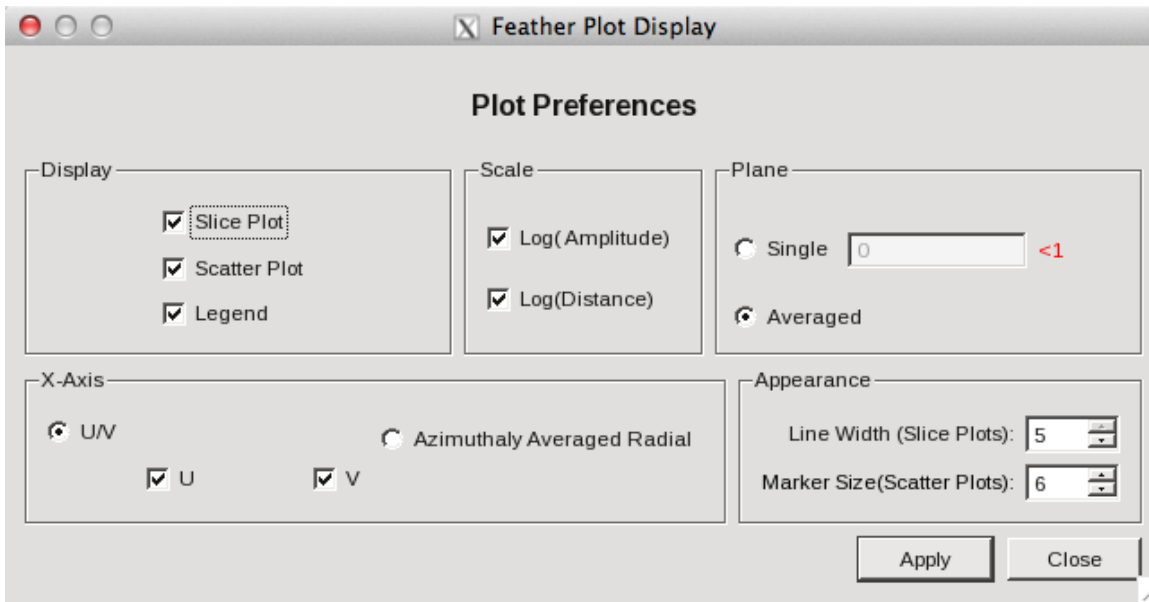


Figure 5.9: The casafeather “customize” window.

`boxit` or the `viewer`. Note that `clean` can take simple boxes or box files as direct input to its `mask` parameter, so these tasks are most useful when direct input to `clean` (or use of interactive `clean`) will not suffice.

There are two ways to construct region files or mask images for use in deconvolution. The `boxit` task will find a set of box regions based upon an input image and control parameters.

5.7.1 Making Deconvolution Regions from an Image (boxit)

The `boxit` task creates “cleanbox” deconvolution regions automatically from an image. It searches the image to find “islands”: all contiguous sets of pixels above the given threshold. The extreme x- and y-pixels of the island are used to determine the corners of a rectangular box that covers each island. The set of boxes are written out into a single region file with extension `.rgn`. `Boxit` works on single-plane images as well as multi-channel images: in the latter case, the thresholding and boxing is done separately in each plane of the image. The output region file from `boxit` can be used as the `mask` input parameter for the `clean` task (§ 5.3).

The parameter inputs for `boxit` are:

```
# boxit :: Box regions in image above given threshold value.
imagenam  =      '' # Name of image to threshold
regionfil  =      '' # Output region file
threshold  =  '0.0mJy' # Threshold value. Must include units.
minsize    =      2 # Minimum number of pixels for a boxable island
diag       =     False # Count diagonal connections?
boxstretch =      1 # Increase box sizes by this many pixels beyond thresholded pixels.
overwrite  =     False # Overwrite existing region file?
```

The `regionfile` parameter specifies the root name of the region file. It will automatically be given `.rgn` as the file extension. The `minsize` parameter specifies the smallest island that qualifies to be boxed. It refers to the total number of pixels in the island. To include pixels connected only on the diagonal as being part of the same island, set the `diag` parameter to `True`. The `boxstretch` parameter increases the size of the boxes beyond the extent of the island, and can range from -1 to 5. For a value of 1 (the default), the box is stretched by one pixel in each outward direction; therefore, each side of the box lengthens by two pixels. Finally, the parameter `overwrite` specifies whether an existing region file can be overwritten.

ALERT: The `boxit` task is a prototype under active development and coded in Python. Eventually we will add functionality to deal with the creation of non-rectangular regions and with multi-plane masks, as well as efficiency improvements.

5.8 Insert an Image Model (ft)

The `ft` task will add a source model (units should be Jy/pixel) or a clean component list to a Measurement Set. This is especially useful if you have a resolved calibrator and you want to start with a model of the source before you derive accurate gain solutions. This is also helpful for self-calibration (see § 5.10 below).

The inputs for `ft` are:

```
# ft :: Insert a source model a visibility set:
vis        =      '' # Name of input visibility file (MS)
field      =      '' # Field selection
```

```

spw           =      'all'      # Spw selection
model        =      ''         # Name of input model image(s)
nterms       =      1          # Number of terms used to model the sky frequency
                                     # dependence
complist     =      ''         # Name of component list
incremental  =      False      # Add to the existing model visibility?
usescratch   =      False      # If True predicted visibility is stored in MODEL_DATA
                                     # column

```

An example on how to do this:

```

ft(vis='n75.ms',           # Start with the visibility dataset n75.ms
   field='1328',          # Select field name '1328+307' (minimum match)
   model='1328.model.image') # Name of the model image you have already

```

This example will add the source model '1328.model.imag' to all entries that match the field name '1328'. If the parameter `usescratch` is set to 'True', `ft` will Fourier transform the source model and fill the `MODEL_DATA` column with the data. This, however, is only needed in special applications and `usescratch=F` is the default.

Alternatively, one can add a clean component list to be used as a model to the MS. The following procedure is an example:

```

# for a point source with no spectral index
cl.addcomponent(flux=0.39, fluxunit='Jy', shape='point', dir='J2000 19h33m09s 15d01m20s')

# for a Gaussian with a spectral index
cl.addcomponent(flux=1.25, fluxunit='mJy', polarization='Stokes',
dir='J2000 19h30m00s 15d00m00s', shape='gaussian', majoraxis='10arcsec',
minoraxis='6arcsec', positionangle='0deg', freq='1.25GHz',
spectrumtype='spectral index', index=-0.8)
###you can add more components if you wish by calling addcomponent repeatedly with different params

##save it to disk
cl.rename('my_component.cl')
cl.close()

## write the model into the measurement set ('mym')
ft(vis='mym', complist='my_component.cl')

```

5.9 Image-plane deconvolution (deconvolve)

If you have only an image (obtained from some telescope) and an image of its point spread function, then you can attempt a simple image-plane deconvolution. Note that for interferometer data, full uv-plane deconvolution using `clean` or similar algorithm is superior!

The default inputs for `deconvolve` are:


```
# deconvolve :: Deconvolving a point spread function from an image

imagename = '' # Name of image to deconvolve
model     = '' # Name of output image to which deconvolved components are stored
psf       = '' # Name of psf or gaussian parameters if psf is assumed gaussian
alg       = 'clark' # Deconvolution algorithm to use
niter     = 10 # number of iteration to use in deconvolution process
gain      = 0.1 # CLEAN gain parameter
threshold = '0.0Jy' # level below which sources will not be deconvolved
mask      = '' # Name of image that has mask to limit region of deconvolution
```

The algorithm (`alg`) options are: `'clark'`, `'hogbom'`, `'multiscale'` or `'mem'`. The `'multiscale'` and `'mem'` options will open the usual set of sub-parameters for these methods.

5.10 Self-Calibration

Once you have a model image or set of model components reconstructed from your data using one of the deconvolution techniques described above, you can use it to refine your calibration. This is called *self-calibration* as it uses the data to determine its own calibration (rather than observations of special calibration sources).

In principle, self-calibration is no different than the calibration process we described earlier (§ 4). In effect, you alternate between calibration and imaging cycles, refining the calibration and the model as you go. The trick is you have to be careful, as defects in early stages of the calibration can get into the model, and thus prevent the calibration from improving. In practice, it is best to not clean very deeply early on, so that the CLEAN model contains correct components only.

One important thing to keep in mind is that the self-calibration relies upon having the most recent source model inside the MS. This is indeed the case if you follow the imaging (using `clean`) directly by the self-calibration. If you have done something strange in between and have lost or overwritten source model (for example done some extra cleaning that you do not want to keep), then use the `ft` task (see § 5.8 above), which adds a source model image or clean component lists to an MS.

Likewise, during self-calibration (once you have a new calibration solution) the imaging part relies upon having the `CORRECTED_DATA` column contain the self-calibrated data. This is done with the `applycal` task (§ 4.6.1).

The `clearcal` command can be used during the self-calibration if you need to clear the `CORRECTED_DATA` column and revert to the original `DATA`. If you need to restore the `CORRECTED_DATA` to any previous stage in the self-calibration, use `applycal` again with the appropriate calibration tables.

ALERT: In later patches we will change the tasks so that users need not worry what is contained in the MS scratch columns and how to fill them. CASA will handle that underneath for you!

For now, we refer the user back to the calibration chapter for a reminder on how to run the calibration tasks.

5.11 Parallelized Cleaning

Parallelized cleaning on multi-core systems will be integrated in the new task `tclean` which is currently **experimental**. This task will replace older code, e.g., `pclean` which should not be used at this stage.

5.12 Examples of Imaging

The data reduction tutorials on casaguides.nrao.edu provide walkthroughs for high and low frequency, spectral line and polarization imaging techniques.

Chapter 6

Image Analysis

Once data has been calibrated (and imaged in the case of synthesis data), the resulting image or image cube must be displayed or analyzed in order to extract quantitative information, such as statistics or moment images. In addition, there need to be facilities for the coordinate conversion of images for direct comparison.

The image analysis tasks are:

- `imhead` — summarize and manipulate the “header” information in a CASA image (§ 6.2)
- `imsubimage` — Create a (sub)image from a region of the image (§ 6.3)
- `imcontsub` — perform continuum subtraction on a spectral-line image cube (§ 6.4)
- `imfit` — image plane Gaussian component fitting (§ 6.5)
- `immath` — perform mathematical operations on or between images (§ 6.6)
- `immoments` — compute the moments of an image cube (§ 6.7)
- `impv` — generate a position-velocity diagram along a slit (§ 6.8)
- `imstat` — calculate statistics on an image or part of an image (§ 6.9)

Inside the Toolkit:

Image analysis is handled in the `ia` tool. Many functions exist there, including region statistics and image math. See § 6.25 below for more information.

- **imval** — extract the data and mask values from a pixel or region of an image (§ 6.10)
- **imtrans** — reorder the axes of an image or cube (§ 6.11)
- **imcollapse** — collapse image along one or more axes by aggregating pixel values along that axis (§ 6.12)
- **imregrid** — regrid an image onto the coordinate system of another image (§ 6.13)
- **imreframe** — change the frame in which the image reports its spectral values (§ 6.14)
- **imrebin** — rebin an image (§ 6.15)
- **specsmooth** — 1-dimensional smooth images in the spectral and angular directions (§ 6.16)
- **imsmooth** — 2-dimensional smooth images in the spectral and angular directions (§ 6.17)
- **specfit** — fit 1-dimensional Gaussians, polynomial, and/or Lorentzians models to an image or image region (§ 6.18)
- **rmfit** — Calculation of rotation measures (§ 6.19)
- **spxfit** — Calculation of Spectral Indices and higher order polynomials (§ 6.20)
- **makemask** — image mask handling (§ 6.21)
- **slsearch** — query a subset of the Splatalogue spectral line catalog (§ 6.22)
- **splattotable** — convert a file exported from Splatalogue to a CASA table (§ 6.23)
- **importfits** — import a FITS image into a CASA *image* format table (§ 6.24.2)
- **exportfits** — write out an image in FITS format (§ 6.24.1)

There are other tasks which are useful during image analysis. These include:

- **viewer** — there are useful region statistics and image cube slice and profile capabilities in the viewer (§ 7)

We also give some examples of using the CASA Toolkit to aid in image analysis (§ 6.25).

6.1 Common Image Analysis Task Parameters

We now describe some sets of parameters are common to the image analysis. These should behave the same way in any of the tasks described in this section that they are found in.

ALERT: As of CASA 4.0.0 we introduced changed `clean` such that for large fractional bandwidths, if the PSF changes more than half a pixel, a different beam will be used for each image plane. These cubes are supported in all image analysis tasks.

6.1.1 Input Image (`imagename`)

The input image typically is an image cube. Most analysis tasks and tools also accept complex valued images.

6.1.2 Region Selection (`box`)

Direction (e.g. RA, Dec) areal selection in the image analysis tasks is controlled by the `box` parameter or through the `region` parameter (§ 6.1.6). Note that one should either specify a region (recommended) or any of `box/chans/stokes`. Specifying both at the same time will give priority to the command line inputs in `'chans'` and `'stokes'` but will keep the region file specification for the spatial region selection.

The `box` parameter selects spatial rectangular areas:

```

box      =      ''      # Select one or more box regions

#      string containing blcx,blcy,trcx,trcy

#      A box selection in the directional portion of an image.
#      The directional portion of an image are the axes for right
#      ascension and declination, for example. Boxes are specified
#      by their bottom-left corner (blc) and top-right corner (trc)
#      as follows: blcx, blcy, trcx, trcy;
#      ONLY pixel values acceptable at this time.
#      Default: none (all);
#      Example: box='0,0,50,50'
```

To get help on `box`, see the in-line help

```
help(par.box)
```

6.1.3 Plane Selection (chans, stokes)

The channel, frequency, or velocity plane(s) of the image is chosen using the `chans` parameter:

```
chans      =      '' # Select the channel(spectral) range

#          string containing channel range

#          immath, imstat, and imcontsub - takes a string listing
#          of channel numbers, velocity, and/or frequency
#          numbers, much like the spw parameter
#          Only channel numbers acceptable at this time.
#          Default: none (all);
#          Example: chans='3~20'
#                   chans="0,3,4,8"
#                   chans="3~20,50,51"
```

`chans` can also be set in the CASA region format to allow settings in frequency and velocity, e.g.

```
chans=("range=[-50km/s,50km/s], restfreq=100GHz, frame=LSRK")
```

this example would even define a new velocity system independent of the one in the image itself. If the rest frequency and velocity frame within the image are being used, the latter two entries are not needed. The parentheses are needed when the call is in a single command.

A frequency selection looks as follows:

```
chans=("range=[100GHz,100.125GHz] ")
```

The polarization plane(s) of the image is chosen with the `stokes` parameter:

```
stokes     =      '' # Stokes params to image (I,IV,IQU,IQUV)

#          string containing Stokes selections

#          Stokes parameters to image, may or may not be separated
#          by commas but best if you use commas.
#          Default: none (all); Example: stokes='IQUV';
#          Example:stokes='I,Q'
#          Options: 'I','Q','U','V',
#                   'RR','RL','LR','LL',
#                   'XX','YX','XY','YY',...
```

To get help on these parameters, see the in-line help

```
help(par.chans)
help(par.stokes)
```

Sometimes, as in the `immoments` task, the channel/plane selection is generalized to work on more than one axis type. In this case, the `planes` parameter is used. This behaves like `chans` in syntax.

6.1.4 Lattice Expressions (expr)

Lattice expressions are strings that describe operations on a set of input images to form an output image. These strings use the *Lattice Expression Language* (LEL). LEL syntax is described in detail in AIPS++ Note 223

http://casa.nrao.edu/aips2_docs/notes/223/index.shtml

The `expr` string contains the LEL expression:

```
expr      =      ' ' # Mathematical expression using images
#
#      string containing LEL expression
#
#      A mathematical expression, with image file names.
#      image file names must be enclosed in double quotes (")
#      Default: none
#      Example: expr='min("image2.im")+(2*max("image1.im"))'
#
#      Available functions in the expr and mask parameters:
#      pi(), e(), sin(), sinh(), asinh(), cos(), cosh(), tan(), tanh(),
#      atan(), exp(), log(), log10(), pow(), sqrt(), complex(), conj()
#      real(), imag(), abs(), arg(), phase(), amplitude(), min(), max()
#      round(), isgn(), floor(), ceil(), rebin(), spectralindex(), pa(),
#      iif(), indexin(), replace(), ...
```

For examples using LEL `expr`, see § 6.6.1 below. Note that in `immath`, shortcut names have been given to the images provided by the user in `imagename` that can be used in the LEL expression, for the above example:

```
imagename=['image2.im','image1.im']
expr='min(IM0)+(2*max(IM1))'
```

ALERT: LEL expressions use 0-based indices. Also, the functions must be lowercase (in almost all cases we know about). Numbers in filenames may be interpreted as such and not strings. Some special characters may also need to be escaped. It is advisable to use double quotes outside single quotes to make such strings being accepted as filenames instead of functions, e.g. `"5sigma+file.im"`.

6.1.5 Masks (mask)

A `mask` can be used to define whether part of an image is used or not. There are different options for masks:

- an image cube with Boolean True/False values
- an image cube with zero and non-zero values

- an LEL string for a condition.

Using image cubes is useful to mask on a pixel by pixel basis, where False and zeros mark masked pixels. Both versions can be converted into each other `makemask` (§ 6.21). Some analysis task show an optional `stretch` parameter which is useful, e.g. to expand a single plane mask to an entire cube along the spectral axis.

An LEL string (see § 6.1.4 above) can be an on-the-fly (OTF) mask expression or refer to an image pixel mask.

```
mask      =      '' # Mask to be applied to the images

#          string containing LEL expression

#          Name of mask applied to each image in the calculation
#          Default '' means no mask;
#          Example: mask='ngc5921.clean.cleanbox.mask">0.5'
#          mask='mask(ngc5921.clean.cleanbox.mask)'
```

Note that the mask file supplied in the `mask` parameter must have the same shape, same number of axes and same axes length, as the images supplied in the `expr` parameter, with one exception. The mask may be missing some of the axes — if this is the case then the mask will be expanded along these axes to become the same shape.

For examples using `mask`, see § 6.6.2 below.

6.1.6 Regions (region)

The `region` parameter points to a CASA region which can be directly specified or listed in a `ImageRegion` file. An `ImageRegion` file can be created with the CASA `viewer`'s region manager (§ 7.4.3). Or directly using the CASA region syntax (Chapter D). Typically `ImageRegion` files will have the suffix `'.crtf'` for CASA Region Text Format.

Alert: When both the region parameter and any of `box/chans/stokes` are specified simultaneously, the task may perform unwanted selections. Only specify one of these (sets of) parameters. We recommend the use of CASA regions and may remove the `box/chans/stokes` selection in later releases.

For example:

```
region='circle[[18h12m24s, -23d11m00s], 2.3arcsec]'
```

or

```
region='myimage.im.crtf'
```


for to specify a region file.

For the most part, the region parameter in tasks only accepts strings (e.g. file names, region shape descriptions) while the region parameter in ia tool methods only accepts python region dictionaries (e.g. produced using the rg tool).

6.2 Image Header Manipulation (imhead)

To summarize and change keywords and values in the “header” of your image, use the `imhead` task. Its inputs are:

```
# imhead :: List, get and put image header parameters
imagername      =      ''      # Name of the input image
mode            = 'summary'    # imhead options: add, del,
                                # get, history, list, put, summary
                                # Give a full listing of
                                # beams or just a short summary?
                                # Only used when the image has multiple beams
                                # and mode="summary".
verbose         =      False
```

The `mode` parameter controls the operation of `imhead`.

Setting `mode='summary'` will print out a summary of the image properties and the header to the logger.

Setting `mode='list'` prints out a list of the header keywords and values to the terminal.

The `mode='get'` allows the user to retrieve the current value for a specified keyword `hdkey`:

```
mode            =      'get'    # imhead options: list, summary, get, put
hdkey           =      ''      # The FITS keyword
```

Note that to catch this value, you need to assign it to a Python variable. See § 1.4.3 for more on return values.

The `mode='put'` allows the user to replace the current value for a given keyword `hditem` with that specified in `hdvalue`. There are two sub-parameters that are opened by this option:

```
mode           =      'put'    # imhead options: list, summary, get, put
hdkey          =      ''      # The FITS keyword
hdvalue        =      ''      # Value of hdkey
hdtype         =      ''      # Data type of the header keyword.
hdcomment      =      ''      # Comment associated with the header keyword
```

WARNING: Be careful when using `mode='put'`. This task does no checking on whether the values you specify (e.g. for the axes types) are valid, and you can render your image invalid. Make sure you know what you are doing when using this option!

6.2.1 Examples for imhead

Here is an example – we can print the summary to the logger:

```
CASA <51>: imhead('ngc5921.demo.cleanimg.image',mode='summary')
```

prints in the logger:

```
##### Begin Task: imhead          #####
Image name       : ngc5921.demo.cleanimg.image
Object name      : N5921_2
Image type       : PagedImage
Image quantity   : Intensity
Pixel mask(s)    : None
Region(s)        : None
Image units      : Jy/beam
Restoring Beam   : 52.3782 arcsec, 45.7319 arcsec, -165.572 deg

Direction reference : J2000
Spectral reference   : LSRK
Velocity type       : RADIO
Rest frequency      : 1.42041e+09 Hz
Pointing center     : 15:22:00.000000 +05.04.00.000000
Telescope           : VLA
Observer           : TEST
Date observation    : 1995/04/13/00:00:00
Telescope position: [-1.60119e+06m, -5.04198e+06m, 3.55488e+06m] (ITRF)

Axis Coord Type      Name                Proj Shape Tile   Coord value at pixel   Coord incr Units
-----
0  0  Direction Right Ascension  SIN  256  64  15:22:00.000  128.00 -1.500000e+01 arcsec
1  0  Direction Declination      SIN  256  64 +05.04.00.000  128.00  1.500000e+01 arcsec
2  1  Stokes Stokes                1    1                I
3  2  Spectral Frequency         46   8  1.41279e+09   0.00  2.4414062e+04 Hz
      Velocity                 1607.99  0.00 -5.152860e+00 km/s

##### End Task: imhead
```

If the beam size per plane differs, the beam information will be displayed for the channel with the smallest beam, the one with the largest beam, and the channel closest to the median beam size. E.g.,

```
Restoring Beams
Pol  Type Chan   Freq   Vel
I    Max   0 9.680e+08  0 39.59 arcsec x 22.77 arcsec pa=-70.57 deg
I    Min 511 1.990e+09 -316516 20.36 arcsec x 12.05 arcsec pa=-65.67 deg
I    Median 255 1.478e+09 -157949 27.11 arcsec x 15.54 arcsec pa=-70.36 deg
```

If `verbose=T` the beam information for every plane will be provided.

If you choose `mode='list'`, you get the summary in the logger and a listing of keywords and values to the terminal:

```

CASA <52>: imhead('ngc5921.demo.cleaning.image',mode='list')
Out[52]:
{'beammajor': 52.378242492675781,
 'beamminor': 45.731891632080078,
 'beampa': -165.5721435546875,
 'bunit': 'Jy/beam',
 'cdelt1': '-7.27220521664e-05',
 'cdelt2': '7.27220521664e-05',
 'cdelt3': '1.0',
 'cdelt4': '24414.0625',
 'crpix1': 128.0,
 'crpix2': 128.0,
 'crpix3': 0.0,
 'crpix4': 0.0,
 'crval1': '4.02298392585',
 'crval2': '0.0884300154344',
 'crval3': 'I',
 'crval4': '1412787144.08',
 'ctype1': 'Right Ascension',
 'ctype2': 'Declination',
 'ctype3': 'Stokes',
 'ctype4': 'Frequency',
 'cunit1': 'rad',
 'cunit2': 'rad',
 'cunit3': '',
 'cunit4': 'Hz',
 'datamax': ' Not Known ',
 'datamin': -0.010392956435680389,
 'date-obs': '1995/04/13/00:00:00',
 'equinox': 'J2000',
 'imtype': 'Intensity',
 'masks': ' Not Known ',
 'maxpixpos': array([134, 134,  0,  38], dtype=int32),
 'maxpos': '15:21:53.976, +05.05.29.998, I, 1.41371e+09Hz',
 'minpixpos': array([117,  0,  0,  21], dtype=int32),
 'minpos': '15:22:11.035, +04.31.59.966, I, 1.4133e+09Hz',
 'object': 'N5921_2',
 'observer': 'TEST',
 'projection': 'SIN',
 'reffreqtype': 'LSRK',
 'restfreq': [1420405752.0],
 'telescope': 'VLA'}

```

Note that this list is a return value and can be captured in a variable:

```
mylist = imhead('ngc5921.demo.cleaning.image',mode='list')
```

The values for these keywords can be queried using `mode='get'`. At this point you should capture the return value:

```
CASA <53>: mybmaj = imhead('ngc5921.demo.cleaning.image',mode='get',hdkey='beammajor')
```

```
CASA <54>: mybmaj
```

```

Out[54]: {'unit': 'arcsec', 'value': 52.378242492699997}

CASA <55>: myobserver = imhead('ngc5921.demo.cleanimg.image',mode='get',hdkey='observer')

CASA <56>: print myobserver
{'value': 'TEST', 'unit': ''}

```

You can set the values for these keywords using `mode='put'`. For example:

```

CASA <57>: imhead('ngc5921.demo.cleanimg.image',mode='put',hdkey='observer',hdvalue='CASA')
Out[57]: 'CASA'

CASA <58>: imhead('ngc5921.demo.cleanimg.image',mode='get',hdkey='observer')
Out[58]: {'unit': '', 'value': 'CASA'}

```

6.3 Extracting sub-images (imsubimage)

The task `imsubimage` provides a way to extract a smaller data cube from a bigger one. The inputs are:

```

# imsubimage :: Create a (sub)image from a region of the image
imasename      = ''          # Input image name. Default is unset.
outfile        = ''          # Output image name. Default is unset.
region         = ''          # Region of interest. Default is whole image.
mask           = ''          # Mask to use. See help par.mask. Default is
                             # none.
dropdeg        = False       # Drop degenerate axes
overwrite      = False       # Overwrite (unprompted) pre-existing output
                             # file?
verbose        = True        # Post additional informative messages to the
                             # logger
stretch        = False       # Stretch the mask if necessary and possible?
wantreturn     = True        # If True, return an image analysis tool
                             # associated with the subimage. If False,
                             # return nothing.

```

The `region` keyword defines the size of the smaller cube and is specified via the CASA region CRTF syntax. E.g.

```
region='box [ [ 100pix , 130pix ] , [120pix, 150pix ] ]'
```

will extract the portion of the image that is between pixel coordinates (100,130) and (12,150). `dropdeg=T` is useful to remove axes in the data cube that are degenerate, i.e. axes with a single plane only. A single Stokes I axis is a common example.

6.4 Continuum Subtraction on an Image Cube (imcontsub)

One method to separate line and continuum emission in an image cube is to specify a number of line-free channels in that cube, make a linear fit to the visibilities in those channels, and subtract the fit from the whole cube. Note that the task `uvcontsub` serves a similar purpose; see § 4.7.6 for a synopsis of the pros and cons of either method.

The `imcontsub` task will subtract a polynomial baseline fit to the specified channels from an image cube.

The default inputs are:

```
# imcontsub :: Continuum subtraction on images
imagername = '' # Name of the input image
linefile = '' # Output line image file name
contfile = '' # Output continuum image file name
fitorder = 0 # Polynomial order for the continuum estimation
region = '' # Image region or name to process see viewer
box = '' # Select one or more box regions
chans = '' # Select the channel(spectral) range
stokes = '' # Stokes params to image (I,IV,IQU,IQUV)
```

Area selection using `box` and `region` is detailed in § 6.1.2 and § 6.1.6 respectively.

Image cube plane selection using `chans` and `stokes` are described in § 6.1.3.

ALERT: `imcontsub` has issues when the image does not contain a spectral or stokes axis. Errors are generated when run on an image missing one or both of these axes. You will need to use the Toolkit (e.g. the `ia.adddegaxes` method) to add degenerate missing axes to the image.

6.4.1 Examples for imcontsub)

For example, we first make a clean image without the uv-plane continuum subtraction:

```
# First, run clearcal to clear the uvcontsub results from the
# corrected column
clearcal('ngc5921.demo.src.split.ms')

# Now clean, keeping all the channels except first and last
default('clean')
vis = 'ngc5921.demo.src.split.ms'
imagername = 'ngc5921.demo.nouvcontsub'
mode = 'channel'
nchan = 61
start = 1
width = 1
imsize = [256,256]
psfmode = 'clark'
imagermode = ''
```

```

cell = [15.,15.]
niter = 6000
threshold='8.0mJy'
weighting = 'briggs'
robust = 0.5
mask = [108,108,148,148]
interactive=False
clean()

# It will have made the image:
# -----
# ngc5921.demo.nouvcontsub.image

# You can view this image
viewer('ngc5921.demo.nouvcontsub.image')

```

You can clearly see continuum sources in the image which were removed previously in the script by the use of `uvcontsub`. Let's see if `imcontsub` can work as well.

Using the viewer, it looks like channels 0 through 4 and 50 through 60 are line-free. Then:

```

default('imcontsub')
imagenam = 'ngc5921.demo.nouvcontsub.image'
linefile = 'ngc5921.demo.nouvcontsub.lineimage'
contfile = 'ngc5921.demo.nouvcontsub.contimage'
fitorder = 1
chans     = '0~4,50~60'
stokes    = 'I'
imcontsub()

```

This did not do too badly!

6.5 Image-plane Component Fitting (`imfit`)

The inputs are:

```

# imfit :: Fit one or more elliptical Gaussian components on an image region(s)
imagenam = '' # Name of the input image
box      = '' # Specify one or more box regions for the fit.
region   = '' # Region. See help par.region for specs.
chans    = '' # Spectral channels on which to perform fit. See "help
              # par.chans" for examples.
stokes   = '' # Stokes parameter to fit. If blank, first stokes plane is
              # used.
mask     = '' # Mask to use. See help par.mask. Default is none.
includepix = [] # Range of pixel values to include for fitting.
excludepix = [] # Range of pixel values to exclude for fitting.
residual = '' # Name of output residual image.

```

```

model          =      ''      # Name of output model image.
estimates      =      ''      # Name of file containing initial estimates of component
                                # parameters.
logfile        =      ''      # Name of file to write fit results.
newestimates   =      ''      # File to write fit results which can be used as initial
                                # estimates for next run.
complist       =      ''      # Name of output component list table.
dooff          =      False    # Also fit a zero level offset? Default is False
rms            =      -1      # RMS to use in calculation of uncertainties. Numeric or
                                # valid quantity (record or string). If numeric, it is
                                # given units of the input image. If quantity, units must
                                # conform to image units. If not positive, the rms of the
                                # residual image, in the region of the fit, is used.
noisefwhm     =      ''      # Noise correlation beam FWHM. If numeric value,
                                # interpreted as pixel widths. If quantity (dictionary,
                                # string), it must have angular units.

```

This task will return (as a Python dictionary) the results of the fit, but the results can also be written into a component list table or a logfile.

Note that to fit more than a single component, you *must* provide starting estimates for each component via the `estimates` file. See ‘`help imfit`’ for more details on this. An noise estimate will be calculated automatically or can be provided through the `rms` and `noisefwhm` keywords.

6.5.1 Examples for `imfit`

The following are some examples using the B1608+656 Tutorial

http://casa.nrao.edu/Doc/Scripts/b1608_demo.py

as an example.

```

# First fit only a single component at a time
# This is OK since the components are well-separated and not blended
# Box around component A
xfit_A_res = imfit('b1608.demo.clean2.image', box='121,121,136,136',
                  newestimates='b1608.demo.clean2.newestimate')

# Now extract the fit part of the return value
xfit_A = xfit_A_res['results']['component0']
#xfit_A
# Out[7]:
#{'flux': {'error': array([ 6.73398035e-05,  0.00000000e+00,  0.00000000e+00,
#                          0.00000000e+00]),
#         'polarisation': 'Stokes',
#         'unit': 'Jy',
#         'value': array([ 0.01753742,  0.          ,  0.          ,  0.          ])},
# 'label': ''}

```

```

# 'shape': {'direction': {'error': {'latitude': {'unit': 'arcsec',
#
#                                     'value': 0.00041154866279462775},
#                                     'longitude': {'unit': 'arcsec',
#
#                                                     'value': 0.00046695916589535109}},
#
#          'm0': {'unit': 'rad', 'value': -2.0541102061078207},
#          'm1': {'unit': 'rad', 'value': 1.1439131060384089},
#          'refer': 'J2000',
#          'type': 'direction'},
#
#          'majoraxis': {'unit': 'arcsec', 'value': 0.29100166137741568},
#          'majoraxiserror': {'unit': 'arcsec',
#
#                             'value': 0.0011186420613222663},
#          'minoraxis': {'unit': 'arcsec', 'value': 0.24738110059830495},
#          'minoraxiserror': {'unit': 'arcsec',
#
#                             'value': 0.0013431999725066338},
#          'positionangle': {'unit': 'deg', 'value': 19.369249322401796},
#          'positionangleerror': {'unit': 'rad',
#
#                                 'value': 0.016663189295782171},
#          'type': 'Gaussian'},
# 'spectrum': {'frequency': {'m0': {'unit': 'GHz', 'value': 1.0},
#
#                                'refer': 'LSRK',
#                                'type': 'frequency'},
#
#              'type': 'Constant'}}

# Now the other components
xfit_B_res = imfit('b1608.demo.clean2.image', box='108,114,120,126',
                  newestimates='b1608.demo.clean2.newestimate', append=True)
xfit_B = xfit_B_res['results']['component0']

xfit_C_res = imfit('b1608.demo.clean2.image', box='108,84,120,96')
xfit_C = xfit_C_res['results']['component0']

xfit_D_res = imfit('b1608.demo.clean2.image', box='144,98,157,110')
xfit_D = xfit_D_res['results']['component0']

print ""
print "Imfit Results:"
print "-----"
print "A Flux = %6.4f Bmaj = %6.4f" % (xfit_A['flux']['value'][0], xfit_A['shape']['majoraxis']['value'])
print "B Flux = %6.4f Bmaj = %6.4f" % (xfit_B['flux']['value'][0], xfit_B['shape']['majoraxis']['value'])
print "C Flux = %6.4f Bmaj = %6.4f" % (xfit_C['flux']['value'][0], xfit_C['shape']['majoraxis']['value'])
print "D Flux = %6.4f Bmaj = %6.4f" % (xfit_D['flux']['value'][0], xfit_D['shape']['majoraxis']['value'])
print ""

```

Now try fitting four components together. For this we will have to provide an estimate file. We will use the clean beam for the estimate of the component sizes:

```

estfile=open('b1608.demo.clean2.estimate','w')
print >>estfile,'# peak, x, y, bmaj, bmin, bpa'
print >>estfile,'0.017, 128, 129, 0.293arcsec, 0.238arcsec, 21.7deg'
print >>estfile,'0.008, 113, 120, 0.293arcsec, 0.238arcsec, 21.7deg'

```



```
print >>estfile,'0.008, 113, 90, 0.293arcsec, 0.238arcsec, 21.7deg'
print >>estfile,'0.002, 151, 104, 0.293arcsec, 0.238arcsec, 21.7deg'
estfile.close()
```

Then, this can be used in `imfit`:

```
xfit_all_res = imfit('b1608.demo.clean2.image',
                    estimates='b1608.demo.clean2.estimate',
                    logfile='b1608.demo.clean2.imfitall.log',
                    newestimates='b1608.demo.clean2.newestimate',
                    box='121,121,136,136,108,114,120,126,108,84,120,96,144,98,157,110')
# Now extract the fit part of the return values
xfit_allA = xfit_all_res['results']['component0']
xfit_allB = xfit_all_res['results']['component1']
xfit_allC = xfit_all_res['results']['component2']
xfit_allD = xfit_all_res['results']['component3']
```

These results are almost identical to those from the individual fits. You can see a nicer printout of the fit results in the logfile.

6.6 Mathematical Operations on an Image (`immath`)

The inputs are:

```
# immath :: Perform math operations on images
imagename      =      ''      # a list of input images
mode           = 'evalexpr'   # mode for math operation (evalexpr, spix, pola, poli)
    expr       =      ''      # Mathematical expression using images
    varnames   =      ''      # a list of variable names to use with the image files

outfile        = 'immath_results.im' # File where the output is saved
mask           =      ''      # Mask to be applied to the images
region         =      ''      # File path which contains an Image Region
box            =      ''      # Select one or more box regions in the input images
chans          =      ''      # Select the channel(spectral) range
stokes         =      'I'     # Stokes params to image (I,IV,IQU,IQUV)
```

In all cases, `outfile` must be supplied with the name of the new output file to create.

The `mode` parameter selects what `immath` is to do.

The default `mode='evalexpr'` lets the user specify a mathematical operation to carry out on one or more input images. The sub-parameter `expr` contains the Lattice Expression Language (LEL) string describing the image operations based on the images in the `imagename` parameter. See § 6.1.4 for more on LEL strings and the `expr` parameter.

Mask specification is done using the `mask` parameter. This can optionally contain an on-the-fly mask expression (in LEL) or point to an image with a pixel mask. See § 6.1.5 for more on the use

of the `mask` parameter. See also § 6.1.4 for more on LEL strings. Sometimes, one would like to use a flat image (e.g. a moment image) mask to be applied to an entire cube. The `stretch=True` subparameter in `mask` allows one to expand the mask to all planes of the cube.

Region selection is carried out through the `region` and `box` parameters. See § 6.1.2 and § 6.1.6 for more on area selection.

Image plane selection is controlled by `chans` and `stokes`. See § 6.1.3 for details on plane selection.

For `mode='evalexpr'`, the standard usage for specifying images to be used in the LEL expression is to provide them as a list in the `imagename` parameter, and then access them in the LEL expression by the names `IM0`, `IM1`, `...`. For example,

```
immath(imagename=['image1.im','image2.im'],expr='IM0-IM1',outfile='ImageDiff.im')
```

would subtract the second image given from the first.

For the special modes `'spix'`, `'pola'`, `'poli'`, the required images for the given operation are to be provided in `imagename` (sometimes in a particular order). **V3.0 ALERT:** For `mode='pola'` you MUST call as a function as in the example below (§ 6.6.1.2), giving the parameters as arguments, or `immath` will fail.

Detailed examples are given below.

6.6.1 Examples for `immath`

In the following, we show a few examples of `immath`. Note that the image names in the `expr` are assumed to refer to existing image files in the current working directory.

6.6.1.1 Simple math

Select a single plane (channel 22) of the 3-D cube and subtract it from the original image:

```
immath(imagename='ngc5921.demo.cleanimg.image',
       expr='IM0',chans='22',
       outfile='ngc5921.demo.chan22.image')
```

Double all values in our image:

```
immath(imagename=['ngc5921.demo.chan22.image'],
       expr='IM0*2.0',
       outfile='ngc5921.demo.chan22double.image' )
```

Square all values in our image:

```
immath(imagename=['ngc5921.demo.chan22.image'],
       expr='IM0^2',
       outfile='ngc5921.demo.chan22squared.image' )
```

Note that the units in the output image are still claimed to be “Jy/beam”, i.e. `immath` will not correctly scale the units in the image for non-linear cases like this. Beware.

You can do other mathematical operations on an image (e.g. trigonometric functions) as well as use scalar results from an image (e.g. max, min, median, mean, variance). You also have access to constants such as `e()` and `pi()` (which are doubles internally, while most images are floats). For example: Take the sine of an image:

```
immath(imagename=['ngc5921.demo.chan22.image', 'ngc5921.demo.chan22squared.image'],
       expr='sin(float(pi())*IM0/sqrt(max(IM1)))',
       outfile='ngc5921.demo.chan22sine.image')
```

Note again that the units are again kept as they were.

Select a single plane (channel 22) of the 3-D cube and subtract it from the original image:

```
immath(imagename='ngc5921.demo.cleanimg.image',
       expr='IM0', chans='22',
       outfile='ngc5921.demo.chan22.image')

immath(imagename=['ngc5921.demo.cleanimg.image', 'ngc5921.demo.chan22.image'],
       expr='IM0-IM1',
       outfile='ngc5921.demo.sub22.image')
```

Note that in this example the 2-D plane gets extended in the third dimension and the 2-D values are applied to each plane in the 3-D cube.

Select and save the inner 1/4 of an image for channels 40,42,44 as well as channels 10 and below:

```
default('immath')
imagename=['ngc5921.demo.cleanimg.image']
expr='IM0'
region='box[[64pix,64pix],[192pix,192pix]]'
chans='<10;40,42,44'
outfile='ngc5921.demo.inner.image'
immath()
```

ALERT: Note that if `chan` selects more than one channel then the output image has a number of channels given by the span from the lowest and highest channel selected in `chan`. In the example above, it will have 45 channels. The ones not selected will be masked in the output cube. If we had set

```
chans = '40,42,44'
```

then there would be 5 output channels corresponding to channels 40,41,42,43,44 of the MS with 41,43 masked. Also, the `chans='<10'` selects channels 0–9.

Note that the `chans` syntax allows the operators '<', '<=', '>', '>='. For example,

```
chans = '<17,>79'
chans = '<=16,>=80'
```

do the same thing.

Divide an image by another, with a threshold on one of the images:

```
immath(imagename=['ngc5921.demo.cleaning.image','ngc5921.demo.chan22.image'],
       expr='IM0/IM1[IM1>0.008]',
       outfile='ngc5921.demo.div22.image')
```

6.6.1.2 Polarization manipulation

The following are some examples using the 3C129 Tutorial

http://casa.nrao.edu/Doc/Scripts/3c129_tutorial.py

as an example.

It is helpful to extract the Stokes planes from the cube into individual images:

```
default('immath')
imagename = '3C129BC.clean.image'
outfile='3C129BC.I'; expr='IM0'; stokes='I'; immath();
outfile='3C129BC.Q'; expr='IM0'; stokes='Q'; immath();
outfile='3C129BC.U'; expr='IM0'; stokes='U'; immath();
outfile='3C129BC.V'; expr='IM0'; stokes='V'; immath();
```

Extract linearly polarized intensity and polarization position angle images:

```
immath(stokes='', outfile='3C129BC.P', mode='poli',
       imagename=['3C129BC.Q','3C129BC.U'], sigma='0.0mJy/beam');
immath(stokes='', outfile='3C129BC.X', mode='pola',
       imagename=['3C129BC.Q','3C129BC.U'], sigma='0.0mJy/beam');
```

V3.0 ALERT: For mode='pola' you MUST call as a function as in this example (giving the parameters as arguments) or `immath` will fail.

Create a fractional linear polarization image:

```
default('immath')
imagename = ['3C129BC.I','3C129BC.Q','3C129BC.U']
outfile='3C129BC.fractional_linpol'
expr='sqrt((IM1^2 + IM2^2)/IM0^2)'
stokes=''
immath()
```

Create a polarized intensity image:

```

default( 'immath')
imagenname = ['3C129BC.Q', '3C129BC.U', '3C129BC.V']
outfile='3C129BC.pol_intensity'
expr='sqrt(IM0^2 + IM1^2 + IM2^2)'
stokes=''
immath()

```

Toolkit Tricks: The following uses the toolkit (§ 6.25). You can make a complex linear polarization ($Q + iU$) image using the `imagepol` tool:

```

# See CASA User Reference Manual:
# http://casa.nrao.edu/docs/casaref/imagepol-Tool.html
#
# Make an imagepol tool and open the clean image
potool = casac.homefinder.find_home_by_name('imagepolHome')
po = potool.create()
po.open('3C129BC.clean.image')
# Use complexlinpol to make a Q+iU image
po.complexlinpol('3C129BC.cmplxlinpol')
po.close()

```

You can now display this in the viewer, in particular overlay this over the intensity raster with the intensity contours. When you load the image, use the LEL:

```
'3C129BC.cmplxlinpol' ['3C129BC.P'>0.0001]
```

which is entered into the LEL box at the bottom of the Load Data menu (§ 7.3.1).

6.6.2 Using masks in `immath`

The `mask` parameter is used inside `immath` to apply a mask to all the images used in `expr` before calculations are done (if you are curious, it uses the `ia.subimage` tool method to make virtual images that are then input in the LEL to the `ia.imagecalc` method).

For example, let's assume that we have made a single channel image using `clean`

```

default('clean')

vis = 'ngc5921.demo.src.split.ms.contsub'
imagenname = 'ngc5921.demo.chan22.cleaning'
mode = 'channel'
nchan = 1
start = 22
step = 1

field = ''
spw = ''
imsize = [256,256]

```

```

cell = [15.,15.]
psfalg = 'clark'
gain = 0.1
niter = 6000
threshold='8.0mJy'
weighting = 'briggs'
rmode = 'norm'
robust = 0.5

mask = [108,108,148,148]

clean()

```

There is now a file 'ngc5921.demo.chan22.cleanimg.mask' that is an image with values 1.0 inside the cleanbox region and 0.0 outside.

We can use this to mask the clean image:

```

default('immath')
imagenname = 'ngc5921.demo.chan22.cleanimg.image'
expr='IMO'
mask='ngc5921.demo.chan22.cleanimg.mask">0.5'
outfile='ngc5921.demo.chan22.cleanimg.imasked'
immath()

```

Toolbox Tricks: Note that there are also *pixel masks* that can be contained in each image. These are Boolean masks, and are implicitly used in the calculation for each image in `expr`. If you want to use the mask in a different image not in `expr`, try it in `mask`:

```

# First make a pixel mask inside ngc5921.demo.chan22.cleanimg.mask
ia.open('ngc5921.demo.chan22.cleanimg.mask')
ia.calcmask('ngc5921.demo.chan22.cleanimg.mask">0.5')
ia.summary()
ia.close()
# There is now a 'mask0' mask in this image as reported by the summary

# Now apply this pixel mask in immath
default('immath')
imagenname='ngc5921.demo.chan22.cleanimg.image'
expr='IMO'
mask='mask(ngc5921.demo.chan22.cleanimg.mask)'
outfile='ngc5921.demo.chan22.cleanimg.imasked1'
immath()

```

Note that nominally the axes of the mask must be congruent to the axes of the images in `expr`. However, one exception is that the image in `mask` can have *fewer* axes (but not axes that exist but are of the wrong lengths). In this case `immath` will extend the missing axes to cover the range in the images in `expr`. Thus, you can apply a mask made from a single channel to a whole cube.

```

# drop degenerate stokes and freq axes from mask image
ia.open('ngc5921.demo.chan22.cleaning.mask')
im2 = ia.subimage(outfile='ngc5921.demo.chan22.cleaning.mymask',dropdeg=True)
im2.summary()
im2.close()
ia.close()
# mymask has only RA and Dec axes

# Now apply this mask to the whole cube
default('immath')
imagename='ngc5921.demo.cleaning.image'
expr='IMO'
mask='"ngc5921.demo.chan22.cleaning.mymask">0.5'
outfile='ngc5921.demo.cleaning.imasked'
immath()

```

For more on masks as used in LEL, see

http://casa.nrao.edu/aips2_docs/notes/223/index.shtml

or in § 6.1.5 above.

6.7 Computing the Moments of an Image Cube (immoments)

For spectral line datasets, the output of the imaging process is an **image cube**, with a frequency or velocity channel axis in addition to the two sky coordinate axes. This can be most easily thought of as a series of **image planes** stacked along the spectral dimension.

A useful product to compute is to collapse the cube into a *moment* image by taking a linear combination of the individual planes:

$$M_m(x_i, y_i) = \sum_k^N w_m(x_i, y_i, v_k) I(x_i, y_i, v_k) \quad (6.1)$$

for pixel i and channel k in the cube I . There are a number of choices to form the m moment, usually approximating some polynomial expansion of the intensity distribution over velocity mean or sum, gradient, dispersion, skew, kurtosis, etc.). There are other possibilities (other than a weighted sum) for calculating the image, such as median filtering, finding minima or maxima along the spectral axis, or absolute mean deviations. And the axis along which to do these calculation need not be the spectral axis (i.e. do moments along Dec for a RA-Velocity image). We will treat all of these as generalized instances of a “moment” map.

The **immoments** task will compute basic moment images from a cube. The default inputs are:

```

# immoments :: Compute moments of an image cube:
imagename = ' ' # Input image name

```

```

moments      =      [0]  # List of moments you would like to compute
axis         = 'spectral' # The moment axis: ra, dec, lat, long, spectral, or stokes
region       =      ''   # Image Region. Use viewer
box          =      ''   # Select one or more box regions
chans        =      ''   # Select the channel(spectral) range
stokes       =      ''   # Stokes params to image (I,IV,IQU,IQUV)
mask         =      ''   # mask used for selecting the area of the
                # image to calculate the moments on
includepix   =      -1   # Range of pixel values to include
excludepix   =      -1   # Range of pixel values to exclude
outfile      =      ''   # Output image file name (or root for multiple moments)

```

This task will operate on the input file given by `imagename` and produce a new image or set of images based on the name given in `outfile`.

The `moments` parameter chooses which moments are calculated. The choices for the operation mode are:

```

moments=-1 - mean value of the spectrum
moments=0  - integrated value of the spectrum
moments=1  - intensity weighted coordinate; traditionally used to get
            'velocity fields'
moments=2  - intensity weighted dispersion of the coordinate; traditionally
            used to get 'velocity dispersion'
moments=3  - median of I
moments=4  - median coordinate
moments=5  - standard deviation about the mean of the spectrum
moments=6  - root mean square of the spectrum
moments=7  - absolute mean deviation of the spectrum
moments=8  - maximum value of the spectrum
moments=9  - coordinate of the maximum value of the spectrum
moments=10 - minimum value of the spectrum
moments=11 - coordinate of the minimum value of the spectrum

```

The meaning of these is described in the CASA Toolkit Manual, that describes the associated `image.moments` tool:

<http://casa.nrao.edu/docs/CasaRef/image.moments.html>

The `axis` parameter sets the axis along which the moment is “collapsed” or calculated. Choices are: `'ra'`, `'dec'`, `'lat'`, `'long'`, `'spectral'`, or `'stokes'`. A standard moment-0 or moment-1 image of a spectral cube would use the default choice `'spectral'`. One could make a position-velocity map by setting `'ra'` or `'dec'`.

The `includepix` and `excludepix` parameters are used to set ranges for the inclusion and exclusion of pixels based on values. For example, `includepix=[0.05,100.0]` will include pixels with values from 50 mJy to 1000 Jy, and `excludepix=[100.0,1000.0]` will exclude pixels with values from 100 to 1000 Jy.

If a single moment is chosen, the `outfile` specifies the exact name of the output image. If multiple `moments` are chosen, then `outfile` will be used as the root of the output filenames, which will get different suffixes for each moment.

For image cubes that contain different beam sizes for each plane, `immoments` will smooth all planes to the largest beam size first, then collapse to the desired moment.

6.7.1 Hints for using (`immoments`)

In order to make an unbiased moment-0 image, do not put in any thresholding using `includepix` or `excludepix`. This is so that the (presumably) zero-mean noise fluctuations in off-line parts of the image cube will cancel out. If your image has large biases, like a pronounced clean bowl due to missing large-scale flux, then your moment-0 image will be biased also. It will be difficult to alleviate this with a threshold, but you can try.

To make a usable moment-1 (or higher) image, on the other hand, it is critical to set a reasonable threshold to exclude noise from being added to the moment maps. Something like a few times the rms noise level in the usable planes seems to work (put into `includepix` or `excludepix` as needed. Also use `chans` to ignore channels with bad data.

6.7.2 Examples using (`immoments`)

Below is an example for `immoments`:

```
default('immoments')
imagenam = 'ngc5921.demo.cleaning'
# Do first and second spectral moments
axis = 'spectral'
chans = ''
moments = [0,1]
# Need to mask out noisy pixels, currently done
# using hard global limits
excludepix = [-100,0.009]
outfile = 'ngc5921.demo.moments'

immoments()

# It will have made the images:
# -----
# ngc5921.demo.moments.integrated
# ngc5921.demo.moments.weighted_coord
```

Other examples of NGC2403 (a moment zero image of a VLA line dataset) and NGC4826 (a moment one image of a BIMA CO line dataset) are shown in Figure 6.1.

ALERT: We are working on improving the thresholding of planes beyond the global cutoffs in `includepix` and `excludepix`.

PV diagrams are generated by placing a “slicing” a datacube through the RA/DEC planes. The “slit” can be defined either by start/end coordinates or by a length, center coordinate, and position angle. Averaged over the width of the ‘slit’ the image cube values are then stored in a new image with position and velocity as the two axes. The slit position is specified by a start and end pixel in the RA/DEC plane of the data cube. An angular unit can be set to define what is stored in the resulting pV image.

6.9 Computing image statistics (imstat)

The `imstat` task will calculate statistics on a region of an image, and return the results as a return value in a Python dictionary.

The inputs are:

```
# imstat :: Displays statistical information from an image or image region
imagenam     =      ''      # Name of the input image
axes         =      -1      # List of axes to evaluate statistics over. Default is
                           # all axes.
region       =      ''      # Image Region or name. Use Viewer
box          =      ''      # Select one or more box regions
chans        =      ''      # Select the channel(spectral) range. See "help
                           # par.chans" for examples.
stokes       =      ''      # Stokes params to image (I,IV,IQU,IQV). Default "" =>
                           # include all
listit       =      True    # Print stats and bounding box to logger?
verbose      =      False   # Print additional messages to logger?
mask         =      ''      # Mask to use. See help par.mask. Default is none.
logfile      =      ''      # Name of file to write fit results.
algorithm    =      'classic' # Algorithm to use. Supported values are "chauenet",
                           # "classic", "fit-half", and "hinges-fences". Minimum
                           # match is supported.
               clmethod     =      'auto' # Method to use for calculating classical statistics.
                           # Supported methods are "auto", "tiled", and
                           # "framework". Ignored if algorithm is not "classic".
```

Area selection using `region` and `mask` is detailed in § 6.1.6 and (§ 6.1.5) respectively.

Plane selection is controlled by `chans` and `stokes`. See § 6.1.3 for details on plane selection.

The parameter `axes` will select the dimensions that the statistics is calculated over. Typical data cubes have axes like: RA axis 0, DEC axis 1, Velocity axis 2. So, e.g. `axes=[0,1]` would be the most common setting to calculate statistics per spectral channel.

A typical output of `imstat` on a cube with `axes=[0,1]` and `algorithm='classic'` (default) looks like:

```
No region specified. Using full positional plane.
Using all spectral channels.
Using polarizations ALL
```

```

Determining stats for image IRC10216_HC3N.cube_r0.5.image
Set region from supplied region record
Statistics calculated using Classic algorithm
Regions ---
  -- bottom-left corner (pixel) [blc]:  [0, 0, 0, 0]
  -- top-right corner (pixel) [trc]:    [299, 299, 0, 63]
  -- bottom-left corner (world) [blcf]: 09:48:01.492, +13.15.40.658, I, 3.63994e+10Hz
  -- top-right corner (world) [trcf]:   09:47:53.299, +13.17.40.258, I, 3.63915e+10Hz
No region specified. Using full positional plane.
Using all spectral channels.
Using polarizations ALL
Selected bounding box :
  [0, 0, 0, 0] to [299, 299, 0, 63] (09:48:01.492, +13.15.40.658, I, 3.63994e+10Hz to 09:47:53.299, +13.17.40.258
#      Frequency  Frequency(Plane)  Npts      Sum      Mean      Rms      Std dev      Minimum
3.63993552e+10      0  9.000000e+04  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
3.63992302e+10      1  9.000000e+04  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
3.63991052e+10      2  9.000000e+04  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
3.63989802e+10      3  9.000000e+04  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
3.63988551e+10      4  9.000000e+04  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
3.63987301e+10      5  9.000000e+04  6.069948e-01  6.744386e-06  1.534640e-03  1.534634e-03  -6.355108e-03
3.63986051e+10      6  9.000000e+04  2.711720e-01  3.013023e-06  1.538071e-03  1.538077e-03  -6.165663e-03
3.63984801e+10      7  9.000000e+04  2.501259e-01  2.779177e-06  1.578049e-03  1.578056e-03  -6.771976e-03
3.63983551e+10      8  9.000000e+04 -3.706732e-01 -4.118591e-06  1.607191e-03  1.607194e-03  -8.871284e-03

```

where the header information provides the specifications of the data that were selected followed by the table with the frequency values of the lanes, the plane numbers, Npts the number of pixels per plane, and the Sum, Median, RMS, Standard deviations, Minimum, and Maximum of the pixel values for each plane. Similar output is provided when the data is averaged over different axes. The logger output can also be written into or appended to a log file for further processing elsewhere (`logfile` parameter).

Imstat has access to different statistics algorithms. Most of them represent different ways on how to treat distributions that are not Gaussian, in particular to eliminate outlier values from the statistics. Available algorithms are *classic*, where all unmasked pixels are used, *fit-half*, where one (good) half of the distribution is being mirrored across a central value, *hinges-fences*, where the inner quartiles plus a 'fence' data portion is being used, and *chawvenet*, which includes values based on the number of standard deviations from the mean. For more information, see the inline help of the `imstat` task.

6.9.1 Using the task return value

The contents of the return value of `imstat` are in a Python dictionary of key-value sets. For example,

```
xstat = imstat()
```

will assign this to the Python variable `xstat`.

The keys for `xstat` are then:

```

KEYS
blc      - absolute PIXEL coordinate of the bottom left corner of
          the bounding box surrounding the selected region
blcf     - Same as blc, but uses WORLD coordinates instead of pixels
trc      - the absolute PIXEL coordinate of the top right corner

```

of the bounding box surrounding the selected region

```

trcf      - Same as trc, but uses WORLD coordinates instead of pixels
flux      - the integrated flux density if the beam is defined and
           the if brightness units are $Jy/beam$
npts      - the number of unmasked points used
max       - the maximum pixel value
min       - minimum pixel value
maxpos    - absolute PIXEL coordinate of maximum pixel value
maxposf   - Same as maxpos, but uses WORLD coordinates instead of pixels
minpos    - absolute pixel coordinate of minimum pixel value
minposf   - Same as minpos, but uses WORLD coordinates instead of pixels
sum       - the sum of the pixel values:  $\sum I_i$ 
sumsq     - the sum of the squares of the pixel values:  $\sum I_i^2$ 
mean      - the mean of pixel values:
            $\bar{I} = \sum I_i / n$ 
sigma     - the standard deviation about the mean:
            $\sigma^2 = (\sum I_i - \bar{I})^2 / (n-1)$ 
rms       - the root mean square:
            $\sqrt{\sum I_i^2 / n}$ 
median    - the median pixel value (if robust=T)
medabsdevmed - the median of the absolute deviations from the
           median (if robust=T)
quartile  - the inter-quartile range (if robust=T). Find the points
           which are 25% largest and 75% largest (the median is
           50% largest), find their difference and divide that
           difference by 2.

```

For example, an `imstat` call might be

```

default('imstat')
imagenam = 'ngc5921.demo.cleanimg.image' # The NGC5921 image cube
box      = '108,108,148,148'             # 20 pixels around the center
chans    = '21'                          # channel 21

xstat = imstat()

```

In the terminal window, `imstat` reports:

Statistics on `ngc5921.usecase.clean.image`

Region ---

```

-- bottom-left corner (pixel) [blc]: [108, 108, 0, 21]
-- top-right corner (pixel) [trc]:  [148, 148, 0, 21]
-- bottom-left corner (world) [blcf]: 15:22:20.076, +04.58.59.981, I, 1.41332e+09Hz
-- top-right corner( world) [trcf]: 15:21:39.919, +05.08.59.981, I, 1.41332e+09Hz

```

Values --

```

-- flux [flux]:          0.111799236126
-- number of points [npts]: 1681.0
-- maximum value [max]:   0.029451508075

```

```

-- minimum value [min]:      -0.00612453464419
-- position of max value (pixel) [maxpos]:  [124, 131, 0, 21]
-- position of min value (pixel) [minpos]:  [142, 110, 0, 21]
-- position of max value (world) [maxposf]: 15:22:04.016, +05.04.44.999, I, 1.41332e+09Hz
-- position of min value (world) [minposf]: 15:21:45.947, +04.59.29.990, I, 1.41332e+09Hz
-- Sum of pixel values [sum]: 1.32267159822
-- Sum of squared pixel values [sumsq]: 0.0284534543692

```

Statistics ---

```

-- Mean of the pixel values [mean]:      0.000786836167885
-- Standard deviation of the Mean [sigma]: 0.00403944306904
-- Root mean square [rms]:              0.00411418313161
-- Median of the pixel values [median]:   0.000137259965413
-- Median of the deviations [medabsdevmed]: 0.00152346317191
-- Quartile [quartile]:                 0.00305395200849

```

The return value in `xstat` is

CASA <152>: `xstat`

```

Out [152]:
{'blc': array([108, 108,  0, 21]),
 'blcf': '15:22:20.076, +04.58.59.981, I, 1.41332e+09Hz',
 'flux': array([ 0.11179924]),
 'max': array([ 0.02945151]),
 'maxpos': array([124, 131,  0, 21]),
 'maxposf': '15:22:04.016, +05.04.44.999, I, 1.41332e+09Hz',
 'mean': array([ 0.00078684]),
 'medabsdevmed': array([ 0.00152346]),
 'median': array([ 0.00013726]),
 'min': array([-0.00612453]),
 'minpos': array([142, 110,  0, 21]),
 'minposf': '15:21:45.947, +04.59.29.990, I, 1.41332e+09Hz',
 'npts': array([ 1681.]),
 'quartile': array([ 0.00305395]),
 'rms': array([ 0.00411418]),
 'sigma': array([ 0.00403944]),
 'sum': array([ 1.3226716]),
 'sumsq': array([ 0.02845345]),
 'trc': array([148, 148,  0, 21]),
 'trcf': '15:21:39.919, +05.08.59.981, I, 1.41332e+09Hz'}

```

ALERT: The return dictionary currently includes NumPy `array` values, which have to be accessed by an array index to get the array value. To access these dictionary elements, use the standard Python dictionary syntax, e.g.

```
xstat[<key string>][<array index>]
```

For example, to extract the standard deviation as a number

```

mystddev = xstat['sigma'][0]
print 'Sigma = '+str(xstat['sigma'][0])

```

6.9.2 Examples for imstat

The following are some examples using the B1608+656 Tutorial

http://casa.nrao.edu/Doc/Scripts/b1608_demo.py

as an example.

To extract statistics for the final image:

```
xstat = imstat('b1608.demo.clean2.image')
# Printing out some of these
print 'Max   = '+str(xstat['max'][0])
print 'Sigma = '+str(xstat['sigma'][0])
# results:
# Max   = 0.016796965152
# Sigma = 0.00033631979385
```

In a box around the brightest component:

```
xstat_A = imstat('b1608.demo.clean2.image',box='124,125,132,133')
# Printing out some of these
print 'Comp A Max Flux = '+str(xstat_A['max'][0])
print 'Comp A Max X,Y = ('+str(xstat_A['maxpos'][0])+','+str(xstat_A['maxpos'][1])+')'
# results:
# Comp A Max Flux = 0.016796965152
# Comp A Max X,Y = (128,129)
```

6.10 Extracting data from an image (imval)

The `imval` task will extract the values of the data and mask from a specified region of an image and place in the task return value as a Python dictionary.

The inputs are:

```
# imval :: Get the data value(s) and/or mask value in an image.
imasename = '' # Name of the input image
region    = '' # Image Region. Use viewer
box       = '' # Select one or more box regions
chans     = '' # Select the channel(spectral) range
stokes    = '' # Stokes params to image (I,IV,IQU,IQV)
```

Area selection using `box` and `region` is detailed in § 6.1.2 and § 6.1.6 respectively. By default, `box=''` will extract the image information at the reference pixel on the direction axes.

Plane selection is controlled by `chans` and `stokes`. See § 6.1.3 for details on plane selection. By default, `chans=''` and `stokes=''` will extract the image information in all channels and Stokes planes.

For instance,

```
xval = imval('myimage', box='144,144', stokes='I' )
```

will extract the Stokes I value or spectrum at pixel 144,144, while

```
xval = imval('myimage', box='134,134.154,154', stokes='I' )
```

will extract a 21 by 21 pixel region.

Extractions are returned in NumPy arrays in the return value dictionary, plus some extra elements describing the axes and selection:

```
CASA <2>: xval = imval('ngc5921.demo.moments.integrated')
```

```
CASA <3>: xval
```

```
Out [3]:
{'axes': [[0, 'Right Ascension'],
          [1, 'Declination'],
          [3, 'Frequency'],
          [2, 'Stokes']],
 'blc': [128, 128, 0, 0],
 'data': array([ 0.89667124]),
 'mask': array([ True], dtype=bool),
 'trc': [128, 128, 0, 0],
 'unit': 'Jy/beam.km/s'}
```

extracts the reference pixel value in this 1-plane image. Note that the 'data' and 'mask' elements are NumPy arrays, not Python lists.

To extract a spectrum from a cube:

```
CASA <8>: xval = imval('ngc5921.demo.clean.image',box='125,125')
```

```
CASA <9>: xval
```

```
Out [9]:
{'axes': [[0, 'Right Ascension'],
          [1, 'Declination'],
          [3, 'Frequency'],
          [2, 'Stokes']],
 'blc': [125, 125, 0, 0],
 'data': array([ 8.45717848e-04,  1.93370355e-03,  1.53750915e-03,
                2.88399984e-03,  2.38683447e-03,  2.89159478e-04,
                3.16268904e-03,  9.93389636e-03,  1.88773088e-02,
                3.01138610e-02,  3.14478502e-02,  4.03211266e-02,
                3.82498614e-02,  3.06552909e-02,  2.80734301e-02,
                1.72479432e-02,  1.20884273e-02,  6.13593217e-03,
                9.04005766e-03,  1.71429547e-03,  5.22095338e-03,
                2.49114982e-03,  5.30831399e-04,  4.80734324e-03,
                1.19265869e-05,  1.29435991e-03,  3.75700940e-04,
                2.34788167e-03,  2.72604497e-03,  1.78467855e-03,
                9.74952069e-04,  2.24676146e-03,  1.82263291e-04,
```



```

    1.98463408e-06,  2.02975096e-03,  9.65532148e-04,
    1.68218743e-03,  2.92119570e-03,  1.29359076e-03,
    -5.11484570e-04,  1.54162932e-03,  4.68662125e-04,
    -8.50282842e-04, -7.91683051e-05,  2.95954203e-04,
    -1.30133145e-03]),
'mask': array([ True,  True,  True,  True,  True,  True,  True,  True,  True,  True,
               True,  True,  True,  True,  True,  True,  True,  True,  True,  True,
               True,  True,  True,  True,  True,  True,  True,  True,  True,
               True,  True,  True,  True,  True,  True,  True,  True,  True,  True], dtype=bool),
'trc': [125, 125, 0, 45],
'unit': 'Jy/beam'}

```

To extract a region from the plane of a cube:

```
CASA <13>: xval = imval('ngc5921.demo.clean.image',box='126,128,130,129',chans='23')
```

```
CASA <14>: xval
Out[14]:
{'axes': [[0, 'Right Ascension'],
          [1, 'Declination'],
          [3, 'Frequency'],
          [2, 'Stokes']],
 'blc': [126, 128, 0, 23],
 'data': array([[ 0.00938627,  0.01487772],
                [ 0.00955847,  0.01688832],
                [ 0.00696965,  0.01501907],
                [ 0.00460964,  0.01220793],
                [ 0.00358087,  0.00990202]]),
 'mask': array([[ True,  True],
                [ True,  True],
                [ True,  True],
                [ True,  True],
                [ True,  True]]), dtype=bool),
'trc': [130, 129, 0, 23],
'unit': 'Jy/beam'}
```

```
CASA <15>: print xval['data'][0][1]
0.0148777160794
```

In this example, a rectangular box was extracted, and you can see the order in the array and how to address specific elements.

6.11 Reordering the Axes of an Image Cube (imtrans)

Sometimes data cubes can be in axis orders that are not adequate for processing. The CASA task `imtrans` can change the ordering of the axis:

```

# imtrans :: Reorder image axes
imagename      =      ''      # Name of the input image
outfile        =      ''      # Name of output CASA image.
order          =      ''      # New zero-based axes order.
wantreturn     =      True     # Return an image tool referencing the
                                # transposed image

```

The `order` parameter is the most important input here. It is a string of numbers that shows how axes 0, 1, 2, 3, ... are mapped onto the new cube (note that the first axis has the label 0, as typical in python). E.g. `order='1032'` will reorder the input axis 0 to be axis 1 in the output, input axis 1 to be output axis 0, input axis 2 to output axis 3 (the last axis) and input axis 3 to output axis 2. Alternatively, axes can be specified by their names. E.g., to reorder an image with right ascension, declination, and frequency and reverse the first two, `order=['declination', 'right ascension', 'frequency']` will work. The axes names can be found typing `(ia.coordsys().names())`. Minimum match is supported, so that `order=["d", "f", "r"]` will produce the same results.

Axes can simultaneously be transposed and reversed. To reverse an axis, precede it by a "-". For example, `order='-10-32'` will reverse the direction of the first and third axis of the input image (the zeroth and second axes in the output image).

Example:

Swap the stokes and spectral axes in an RA-Dec-Stokes-Frequency image

```

imagename = "myim.im"
outfile = "outim.im"
order = "0132"
imtrans()

```

or

```

outfile = "myim_2.im"
order = 132
imtrans()

```

or

```

outfile = "myim_3.im"
order = ["r", "d", "f", "s"]
imtrans()

```

or

```

outfile = "myim_4.im"
order = ["rig", "declin", "frequ", "stok"]
imtrans()

```

If the `outfile` parameter is empty, only a temporary image is created; no output image is written to disk. The temporary image can be captured in the returned value (assuming `wantreturn` is true).

6.12 Collapsing an Image Along an Axis (`imcollapse`)

`imcollapse` allows to apply an aggregation function along one or more axes of an image. Functions supported are 'max', 'mean', 'median', 'min', 'rms', 'stdev', 'sum', 'variance' (minimum match supported). The relevant axes will then collapse to a single value or plane (i.e. they will result in a degenerate axis). The functions are specified in the `function` parameter of the `imcollapse` inputs:

```
# imcollapse :: Collapse image along one axis, aggregating pixel values along that axis.
imagenam     =      ''      # Name of the input image
function     =      ''      # Function used to compute aggregation
                                # of pixel values.
axes         =      [0]     # Zero-based axis number(s) or minimal
                                # match strings to collapse.
outfile     =      ''      # Name of output CASA image.
box         =      ''      # Optional direction plane box ("blcx,
                                # blcy, trcx trcy").
    region   =      ''      # Name of optional region file to use.

chans       =      ''      # Optional zero-based contiguous
                                # frequency channel specification.
stokes      =      ''      # Optional contiguous stokes planes
                                # specification.
mask        =      ''      # Optional mask to use.
wantreturn  =      True     # Should an image analysis tool
                                # referencing the collapsed image be
                                # returned?
```

`wantreturn=True` returns an image analysis tool containing the newly created collapsed image.

Example:

`myimage.im` is a 512x512x128x4 (ra,dec,freq,stokes; i.e. in the 0-based system, frequency is labeled as axis 2) image and we want to collapse a subimage of it along its spectral axis avoiding the 8 edge channels at each end of the band, computing the mean value of the pixels (resulting image is 256x256x1x4 in size):

```
imcollapse(imagenam="myimage.im", outfile="collapse_spec_mean.im",
           function="mean", axis=2, box="127,127,383,383", chans="8~119")
```

Note that `imcollapse` will not smooth to a common beam for all planes if they differ. If this is desired, run `imsmooth` before `imcollapse`.

6.13 Regridding an Image (`imregrid`)

It is occasionally necessary to regrid an image onto a new coordinate system. The `imregrid` task will regrid one image onto the coordinate system of another, creating an output image. In this task, the user need only specify the names of the input, template, and output images.

The default inputs are:

```

imregrid :: regrid an image onto a template image
imasename      =      ''          # Name of the source image
template       =      'get'       # A dictionary, refcode, or
                                # name of an image that provides the output shape and coordinate
output        =      ''          # Name for the regridded image
asvelocity    =      True        # Regrid spectral axis in velocity space rather than frequency space
axes          =      [-1]        # The pixel axes to regrid. -1 => all.
interpolation  =      'linear'   # The interpolation method. One of "nearest", "linear", "cubic"
decimate      =      10         # Decimation factor for coordinate grid computation
replicate     =      False       # Replicate image rather than regrid?
overwrite     =      False       # Overwrite (unprompted) pre-existing output file?

```

Inside the Toolkit:

More complex coordinate system and image regridding operation can be carried out in the toolkit. The `coordsys (cs)` tool and the `ia.regrid` method are the relevant components.

The output image will have the data in `imasename` regridded onto the coordinate system provided by the `template` parameter. `template` is used universally for a range of ways to define the grid of the output image:

- *a template image*: specify an image name here and the input will be regridded to the same 3-dimensional coordinate system as the one in `template`. Values are filled in as blanks if they do not exist in the input. Note that the input and template images must have the same coordinate structure to begin with (like 3 or 4 axes, with the same ordering)
- *a coordinate system (reference code)*: to convert from one coordinate frame to another one, e.g. from B1950 to J2000, the `template` parameter can be used to specify the output coordinate system. These following recognized keywords are supported: 'J2000', 'B1950', 'B1950_VLA', 'GALACTIC', 'HADEC', 'AZEL', 'AZELSW', 'AZELNE', 'ECLIPTIC', 'MECLIPTIC', 'TECLIPTIC', 'SUPERGAL'
- *'get'*: This option returns a python dictionary in the {'csys': `csys_record`, 'shap': `shape`} format
- *a python dictionary*: In turn, such a dictionary can be used as a template to define the final grid

6.14 Redefining the Velocity System of an Image (`imreframe`)

`imreframe` can be used to change the velocity system of an image. It is not applying a regridding as a change from radio to optical conventions would require, but it will change the labels of the velocity axes.

```
# imreframe :: Change the frame in which the image reports its spectral values
imagename      =      ''      # Name of the input image
output         =      ''      # Name of the output image; '' => modify input image
outframe       =      'lsrk'   # Spectral frame in which the frequency or velocity
                               # values will be reported by default
restfreq       =      ''      # restfrequency to use for velocity values (e.g.
                               # "1.420GHz" for the HI line)
```

`outframe` defines the velocity frame (LSRK, BARY, etc., see § C.2) of the output image and a rest frequency should be specified to relabel the spectral axis in new velocity units.

6.15 Rebin an Image (`imrebin`)

The task `imrebin` allows one to rebin an image in any spatial or spectral direction:

```
# imrebin :: Rebin an image by the specified integer factors
imagename      =      ''      # Name of the input image
outfile        =      ''      # Output image name.
factor         =      []      # Binning factors for each axis
region         =      ''      # The region to rebin. Default is entire image. Do not
                               # specify region and box/chans simultaneously.
    box         =      ''      # Box in directional plane to rebin. Default is to use the
                               # entire directional plane.
    chans       =      ''      # Channels to rebin. See "help par.chans" for examples.
                               # Default is all channels
    stokes      =      ''      # The correlations to include in the output. Default is
                               # all. Stokes planes cannot be rebinned.

mask           =      ''      # Mask to use. See help par.mask. Default is none.
dropdeg        =      False   # Drop degenerate axes?
crop           =      True    # Remove pixels from the end of an axis to be rebinned if
                               # there are not enough to form an integral bin?
```

where *factor* is a list of integers that provides the numbers of pixels to be binned for each axis. The *crop* parameters controls how pixels at the boundaries are treated if the bin values are not multiple integers of the image dimensions.

Example:

```
imrebin(imagename="my.im", outfile="rebinned.im", factor=[1,2,1,4], crop=T)
```

This leaves RA untouched, bins DEC by a factor of 2, leaves Stokes as is, and bins the spectral axis by a factor of 4. If the input image has a spectral axis with a length that is not a multiple of 4, the *crop=T* setting will discard the remaining 1-3 edge pixels.

6.16 1-dimensional Smoothing (specsmooth)

To gain higher signal-to-noise of data cubes, one can smooth the data along one dimension (for 2-dimensional smoothing, see `imsmooth` § 6.16). Typically this is the spectral axis. Hanning and Boxcar smoothing kernels are available in the task `specsmooth`:

```
# specsmooth :: Smooth an image region in one dimension
imasename      =      ''      # Name of the input image
outfile        =      ''      # Output image name.
region         =      ''      # Region selection. See help par.region for possible
                                # specifications. Default: Do not use a region.
                                # Rectangular box in direction coordinate blc, trc.
                                # Default: entire image ("").
        box     =      ''      #
mask           =      ''      # Mask to use. See help par.mask. Default is none..
axis          =      -1      # The profile axis. Default: use the spectral axis if one
                                # exists, axis 0 otherwise (<0).
function       =  'hanning'   # Convolution function. hanning and boxcar are supported
                                # functions. Minimum match is supported.
dmethod        =      'copy'   # Decimation method. "" means no decimation, "copy" and
                                # "mean" are also supported
(minimum match).
```

The parameter `dmethod='copy'` allows one to only keep every `w`'th channel, if the smoothing kernel has a width of `w`. Leaving this parameter empty will return the same size cube as the input and setting it to `'mean'` will average planes using the kernel width.

6.17 2-dimensional Smoothing; Image Convolution (imsmooth)

A data cube can be smoothed across spatial dimensions with `imsmooth`. The inputs are:

```
# imsmooth :: Smooth an image or portion of an image
imasename      =      ''      # Name of the input image. Must be specified.
kernel         =      'gauss'  # Type of kernel to use. Acceptable values are "b", "box",
                                # or "boxcar" for a boxcar kernel, "g", "gauss", or
                                # "gaussian" for a gaussian kernel, "c", "common", or
                                # "commonbeam" to use the common beam of an image with
                                # multiple beams as the gaussian to which to convolve all
                                # the planes, "i" or "image" to use an image as the
                                # kernel.
        beam    =      ''      # Alternate way of describing a Gaussian. If specified,
                                # must be a dictionary with keys "major", "minor", and
                                # "pa" (or "positionangle"). Do not specify beam if
                                # specifying major, minor, and pa. Example: Example:
                                # {"major": "5arcsec", "minor": "2arcsec", "pa": "20deg"}.
```

```

    targetres      =      False      # If gaussian kernel, specified parameters are to be
    # resolution of output image (True) or parameters of
    # gaussian to convolve with input image (False).
    major          =      ''         # Major axis for the kernels. Standard quantity
    # representation. Must be specified for kernel="boxcar".
    # Example: "4arcsec".
    minor          =      ''         # Minor axis. Standard quantity representation. Must be
    # specified for kernel="boxcar". Example: "2arcsec".
    pa             =      ''         # Position angle used only for gaussian kernel. Standard
    # quantity representation. Example: "40deg".

    region         =      ''         # Region selection. See help par.region. Empty string means
    # use box/chans/stokes if supplied, or else entire image.
    box            =      ''         # Rectangular region specification in directional plane. Do
    # not specify region if you specify box.
    chans          =      ''         # Select the spectral channel range. See "help par.chans"
    # for examples. Do not specify region if you specify
    # chans.
    stokes         =      'I'        # Stokes parameters to image (eg, I,IV,IQU,IQUV). Do not
    # specify region if you specify stokes.
    mask           =      []         # Mask to use. See help par.mask. Default is none.
    stretch       =      False      # Stretch the mask if necessary and possible? See help
    # par.stretch

    outfile        =      ''         # Output image name. Must be specified.
    overwrite      =      False      # Overwrite (unprompted) pre-existing output file?

```

where the cube/image `imagename` will be convolved with a kernel defined in the `kernel` keyword. Kernels 'gauss' and 'boxcar' need the major and minor axes sizes as input, the Gaussian kernel smoothing also requires a position angle. By default, the kernel size defines the kernel itself, i.e. the data will be smoothed with this kernel. If the `targetres` parameter for Gaussian kernels is set to 'True', major and minor axes will be those from the `output` resolution, and the kernel will be adjusted for each plane to arrive at the final resolution.

The 'commonbeam' kernel is to be used when the beam shape is different as a function of frequency. This option will then smooth all planes to a single beam, defined by the largest beam in the cube.

With the 'image' kernel, one can specify an image that will serve as the convolution kernel. A scale factor can be applied, which defaults to flux conservation.

Examples:

1) smoothing to a Gaussian kernel 20" by 10"

```
imsmooth( imagename='my.image', kernel='gauss', major='20arcsec', minor='10arcsec',targetres=T)
```

2) Smoothing using pixel coordinates and a boxcar kernel.

```
imsmooth( imagename='new.image', major='20pix', minor='10pix', kernel='boxcar')
```

6.18 Spectral Line fitting with specfit

`specfit` is a powerful task to perform spectral line fits in data cubes. Three types of fitting functions are currently supported, polynomials, Gaussians, and Lorentzians. `specfit` can fit these functions in two ways: over data that were averaged across a region (`multifit=False`) or on a pixel by pixel basis (`multifit=True`).

```
# specfit :: Fit 1-dimensional Gaussians and/or polynomial models to an image or image region
imagename      =      ''      # Name of the input image
box            =      ''      # Rectangular box in direction coordinate
                                # blc, trc. Default: entire image ("").
region         =      ''      # Region of interest. See help par.region
                                # for possible specifications. Default: Do
                                # not use a region.
chans          =      ''      # Channels to use. Channels must be
                                # contiguous. Default: all channels ("").
stokes         =      ''      # Stokes planes to use. Planes must be
                                # contiguous. Default: all stokes ("").
axis           =      -1      # The profile axis. Default: use the
                                # spectral axis if one exists, axis 0
                                # otherwise (<0).
mask           =      ''      # Mask to use. See help par.mask. Default is
                                # none..
poly           =      -1      # Order of polynomial element. Default: do
                                # not fit a polynomial (<0).
estimates      =      ''      # Name of file containing initial estimates.
                                # Default: No initial estimates ("").
    ngauss     =      1      # Number of Gaussian elements. Default: 1.
    pampest    =      ''      # Initial estimate of PCF profile (gaussian
                                # or lorentzian) amplitudes.
    pcenterest =      ''      # Initial estimate PCF profile centers, in
                                # pixels.
    pfwhmest   =      ''      # Initial estimate PCF profile FWHMs, in
                                # pixels.
    pfix       =      ''      # PCF profile parameters to fix during fit.
    pfunc      =      ''      # PCF singlet functions to fit. "gaussian"
                                # or "lorentzian" (minimal match
                                # supported). Unspecified means all
                                # gaussians.

minpts         =      0      # Minimum number of unmasked points
                                # necessary to attempt fit.
multifit       =      True   # If true, fit a profile along the desired
                                # axis at each pixel in the specified
                                # region. If false, average the non-fit
                                # axis pixels and do a single fit to that
                                # average profile. Default False.
    amp        =      ''      # Name of amplitude solution image. Default:
                                # do not write the image ("").
    amperr     =      ''      # Name of amplitude solution error image.
```



```

        center          =      ''      # Default: do not write the image ("").
        centererr      =      ''      # Name of center solution image. Default: do
        # not write the image ("").
        centererr      =      ''      # Name of center solution error image.
        # Default: do not write the image ("").
        fwhm           =      ''      # Name of fwhm solution image. Default: do
        # not write the image ("").
        fwhmerr        =      ''      # Name of fwhm solution error image.
        # Default: do not write the image ("").
        integral       =      ''      # Prefix of name of integral solution image.
        # Name of image will have gaussian
        # component number appended. Default: do
        # not write the image ("").
        integralerr    =      ''      # Prefix of name of integral error solution
        # image. Name of image will have gaussian
        # component number appended. Default: do
        # not write the image ("").

model                =      ''      # Name of model image. Default: do not write
        # the model image ("").
residual             =      ''      # Name of residual image. Default: do not
        # write the residual image ("").
wantreturn           =      True     # Should a record summarizing the results be
        # returned?
logresults           =      True     # Output results to logger?
gmncomps             =      0        # Number of components in each gaussian
        # multiplet to fit
gmampcn              =      ''      # The amplitude ratio constraints for non-
        # reference components to reference
        # component in gaussian multiplets.
gmcentercn           =      ''      # The center offset constraints (in pixels)
        # for non-reference components to reference
        # component in gaussian multiplets.
gmfwhmcon            =      ''      # The FWHM ratio constraints for non-
        # reference components to reference
        # component in gaussian multiplets.
gmampest             =      [0.0]    # Initial estimate of individual gaussian
        # amplitudes in gaussian multiplets.
gmcenterest          =      [0.0]    # Initial estimate of individual gaussian
        # centers in gaussian multiplets, in
        # pixels.
gmfwhmest            =      [0.0]    # Initial estimate of individual gaussian
        # FWHMss in gaussian multiplets, in pixels.
gmfix                =      ''      # Parameters of individual gaussians in
        # gaussian multiplets to fix during fit.
logfile              =      ''      # File in which to log results. Default is
        # not to write a logfile.
goodamprange         =      [0.0]    # Acceptable amplitude solution range. [0.0]
        # => all amplitude solutions are
        # acceptable.
goodcenterrange      =      [0.0]    # Acceptable center solution range in pixels

```

```

# relative to region start. [0.0] => all
# center solutions are acceptable.
goodfwhmrange = [0.0] # Acceptable FWHM solution range in pixels.
# [0.0] => all FWHM solutions are
# acceptable.
sigma = '' # Standard deviation array or image name.

```

6.18.1 Polynomial Fits

Polynomials can be fit by specifying the polynomial order in `poly`. Negative orders will not fit any polynomials.

6.18.2 Lorentzian and Gaussian Fits

Gaussian and Lorentzian fits are very similar, they both require amplitude, center, and FWHM to be fully specified. All of the following discussion is thus valid for both functions. The parameter `pfunc` controls whether Gaussian or Lorentzian functions are to be used. Default is all Gaussians. `pfunc=["L", "G", "G", "L"]` would use Lorentzian, Gaussian, Gaussian, and Lorentzian components in the order they appear in the estimates file (see below).

6.18.2.1 One or more single Gaussian/Lorentzian

For Gaussian and Lorentzian fits, the task will allow multiple components and `specfit` will try to find the best solution. The parameter space, however, is usually not uniform and to avoid local minima in the goodness-of-fit space, one can provide initial start values for the fits. This can be done either through the parameters `pampest`, `pcenterest`, and `pfwhmest` for the amplitudes, center, and FWHM estimates in image coordinates. `pfix` can take parameters that specify fixed fit values. Any combination of the characters 'p' (peak), 'c' (center), and 'f' (fwhm) are permitted, e.g. "fc" will hold the fwhm and the center constant during the fit. Fixed parameters will have no errors associated with them in the solution. Alternatively, a file with initial values can be supplied by the `estimates` parameter (one Gaussian/Lorentzian parameter set per line). The file has the following format:

```
[peak intensity], [center], [fwhm], [optional fixed parameter string]
```

The first three values are required and must be numerical values. The peak intensity must be expressed in map units, while the center and fwhm must be specified in pixels. The fourth value is optional and if present, represents the parameter(s) that should be held constant during the fit (see above).

An example `estimates` file is:

```

# estimates file indicating that two Gaussians should be fit
# first gaussian estimate, peak=40, center at pixel number 10.5,
# fwhm = 5.8 pixels, all parameters allowed to vary during

```

```
# fit
40, 10.5, 5.8
# second Gaussian, peak = 4, center at pixel number 90.2,
# fwhm = 7.2 pixels, hold fwhm constant
4, 90.2, 7.2, f
# end file
```

and the output of a typical execution, e.g.

```
specfit(imagename='IRC10216_HC3N.cube_r0.5.image',
region='specfit.crtf', multifit=F, estimates='', ngauss=2)
```

('specfit.crtf' is a CASA regions file, see Section D)

will be

```
Fit   :
RA      : 09:47:57.49
Dec     : 13.16.46.46
Stokes  : I
Pixel   : [146.002, 164.499, 0.000, *]
Attempted : YES
Converged : YES
Iterations : 28
Results for component 0:
Type    : GAUSSIAN
Peak    : 5.76 +/- 0.45 mJy/beam
Center  : -15.96 +/- 0.32 km/s
        40.78 +/- 0.31 pixel
FWHM    : 7.70 +/- 0.77 km/s
        7.48 +/- 0.74 pixel
Integral : 47.2 +/- 6.0 mJy/beam.km/s
Results for component 1:
Type    : GAUSSIAN
Peak    : 4.37 +/- 0.33 mJy/beam
Center  : -33.51 +/- 0.58 km/s
        23.73 +/- 0.57 pixel
FWHM    : 15.1 +/- 1.5 km/s
        14.7 +/- 1.5 pixel
Integral : 70.2 +/- 8.8 mJy/beam.km/s
```

If `wantreturn=True` (the default value), the task returns a python dictionary (here captured in a variable with the inventive name of 'fitresults') :

```
fitresults=specfit(imagename='IRC10216_HC3N.cube_r0.5.image', region='specfit.rgn', multifit=F,
estimates='', ngauss=2)
```

The values can then be used by other python code for further processing.

6.18.2.2 Gaussian Multiplets

It is possible to fit a number of Gaussians together, as multiplets with restrictions. All restrictions are relative to a reference Gaussian (the zero'th component of each multiplet). `gncomps` specifies the number of Gaussians for each multiplets, and, in fact, a number of these multiplets can be fit simultaneously. `gncomps=[2,4,3]`, e.g. fits a 2-component Gaussian, a 4-component Gaussian, and a 3-component Gaussian all at once. The initial parameter estimates can be specified with the `gmampest`, `gmcenterest`, and `gmfwhmest` parameters and the estimates are simply listed in the sequence of `gncomps`. E.g. if `gncomps=[2,4,3]` is specified with multiplet G0 consisting of 2 Gaussians a, b, multiplet G1 of 4 Gaussians c, d, e, f, and multiplet G2 of three Gaussians g, h, i, the parameter list in `gm*est` would be like `gm*est=[a,b,c,d,e,f,g,h,i]`.

Restrictions can be specified via the `gmampcon` parameter for the amplitude ratio (non-reference to reference), `gmcentercon` for the offset in pixels (to a reference), and `gmfwhmcon` for the FWHM ratio (non-reference to reference). A value of 0 will not constrain anything. The reference is always the zero'th component in each multiplet, in our example, Gaussians a, c, and g. They cannot be constrained. So `gmncomps=[2, 4, 3]`, `gmampcon= [0 , 0.2, 0 , 0.1, 4.5, 0]`, `gcentercon=[24.2, 45.6, 92.7, 0 , -22.8, -33.5]`, and `gmfwhmcon=""` would constrain Gaussians b relative to a with a 24.2 pixel offset, Gaussian d to c with a amplitude ratio of 0.2 and a 45.6 pixel offset, Gaussian e to c with a offset of 92.7 pixels, etc. Restrictions will overrule any estimates.

The parameters `goodamprange`, `goodcenterrange`, and `goodfwhmrange` can be used to limit the range of amplitude, center, and fwhm solutions for all Gaussians.

6.18.3 Pixel-by-pixel fits

As mentioned above, `specfit` can also fit spectral cubes on a pixel by pixel basis. In this case, one can choose to write none, any or all of the solution and error images for Gaussian/Lorentzian fits via the parameters `amp`, `amperr`, `center`, `centererr`, `fwhm`, and `fwhmerr`. The files written contain all the components along a “component number” axis. They can be inspected by using the “hidden axes” slider in the `viewer` (§7.4.1). Writing analogous images for polynomial coefficients is not yet supported although polynomial fits when `multifit=True` is supported. Best fit coefficients are written to the logger. Pixels for which fits were not attempted or did not converge will be masked as bad.

6.19 Calculation of Rotation Measures `rmfit`

`rmfit` is an image domain task that accepts an input cube image containing Stokes Q and U axes and will generate the rotation measure by performing a least square fit in the image domain to obtain the best fit to the equation $\chi = \chi_0 + RM \times \lambda^2$.

The inputs to `rmfit` are:

```

# rmfit :: Calculate rotation measure.
imagename      =      ''      # Name(s) of the input image(s). Must be specified.
rm             =      ''      # Output rotation measure image name. If not specified, no
                        # image is written.
rmerr          =      ''      # Output rotation measure error image name. If not
                        # specified, no image is written.
pa0            =      ''      # Output position angle (degrees) at zero wavelength image
                        # name. If not specified, no image is written.
pa0err         =      ''      # Output position angle (degrees) at zero wavelength error
                        # image name. If not specified, no image is written.
nturns         =      ''      # Output number of turns image name. If not specified, no
                        # image is written.
chisq          =      ''      # Output reduced chi squared image name. If not specified,
                        # no image is written.
sigma          =      ''      # Estimate of the thermal noise. A value less than 0 means
                        # auto estimate.
rmfg           =      0.0      # Foreground rotation measure in rad/m/m to subtract.
rmmax          =      0.0      # Maximum rotation measure in rad/m/m for which to solve.
                        # IMPORTANT TO SPECIFY.
maxpaerr       =      1e+30    # Maximum input position angle error in degrees to allow in
                        # solution determination.

```

This task generates the rotation measure image from stokes Q and U measurements at several different frequencies. You are required to specify the name of at least one image with a polarization axis containing stokes Q and U planes and with a frequency axis containing more than two pixels. The frequencies do not have to be equally spaced (ie the frequency coordinate can be a tabular coordinate). It will work out the position angle images for you. You may also specify multiple image names, in which case these images will first be concatenated along the spectral axis using `ia.imageconcat()`. The requirements are that for all images, the axis order must be the same and the number of pixels along each axis must be identical, except for the spectral axis which may differ in length between images. The spectral axis need not be contiguous from one image to another.

See also the `imagepol.fourierrotationmeasure` function for a new Fourier-based approach.

Rotation measure algorithms that work robustly are few. The main problem is in trying to account for the $n - \pi$ ambiguity (see Leahy et al, *Astronomy & Astrophysics*, 156, 234 or the MIRIAD manual¹).

The algorithm that this task uses is that of Leahy et al. in their Appendix A.1. But as in all these algorithms, the basic process is that for each spatial pixel, the position angle vs frequency data is fit to determine the rotation measure and the position angle at zero wavelength (and associated errors). An image containing the number of $n - \pi$ turns that were added to the data at each spatial pixel and for which the best fit was found can be written. The reduced χ^2 image for the fits can also be written.

Note that no assessment of curvature (i.e. deviation from the simple linear position angle - λ^2 functional form) is made.

Any combination of output images can be written.

¹<http://www.cfa.harvard.edu/sma/miriad/manuals/SMAguide/smauserhtml/imrm.html>

The parameter *sigma* gives the thermal noise in Stokes Q and U. By default it is determined automatically using the image data. But if it proves to be inaccurate (maybe not many signal-free pixels), it may be specified. This is used for calculating the error in the position angles (via propagation of Gaussian errors).

The argument *maxpaerr* specifies the maximum allowable error in the position angle that is acceptable. The default is an infinite value. From the standard propagation of errors, the error in the linearly polarized position angle is determined from the Stokes Q and U images (at each directional pixel for each frequency). If the position angle error for any pixel exceeds the specified value, the position angle at that pixel is omitted from the fit. The process generates an error for the fit and this is used to compute the errors in the output images.

Note that *maxpaerr* is not used to mask pixels in the output images.

The argument *rmfg* is used to specify a foreground RM value. For example, you may know the mean RM in some direction out of the Galaxy, then including this can improve the algorithm by reducing ambiguity.

The parameter *rmmax* specifies the maximum absolute RM value that should be solved for. This is quite an important parameter. If you leave it at the default, zero, no ambiguity handling will be used. So some a priori information should be supplied; this is the basic problem with rotation measure algorithms.

6.20 Calculation of Spectral Indices and Higher Order Polynomials (spxfit)

This application fits a power logarithmic polynomial or a logarithmic transformed polynomial to pixel values along a specified axis of an image or images. These functions are most often used for fitting the spectral index and higher order terms of a spectrum. A power logarithmic polynomial has the form

$$y = \frac{c_0 x}{D^{(c_1 + c_2 \ln(x/D) + c_3 \ln(x/D)^2 + c_n \ln(x/D)^{(n-1)})}} \quad (6.2)$$

and a logarithmic transformed polynomial is simply the result of this equation after taking the natural log of both sides so that it has the form

$$\ln(y) = c_0 + c_1 \ln(x) + c_2 \ln(x/D)^2 + \dots + c_n \ln(x/D)^n \quad (6.3)$$

Because the logarithm of the ordinate values must be taken before fitting a logarithmic transformed polynomial, all non-positive pixel values are effectively masked for the purposes of fitting.

The coefficients of the two forms are equal to each other except that c_0 in the second equation is equal to $\ln(c_0)$ of the first. In the case of fitting a spectral index, which is traditionally represented as α , is equal to c_1 .

In both cases, D is a normalization constant used so that abscissa values are closer to unity when they are sent to the fitter. This generally improves the probability that the fit will converge. This parameter may be specified via the `div` parameter. A value of 0 (the default) indicates that the application should determine a reasonable value for D , which is determined via

$$D = 10^{\int (\log 10(\sqrt{\min(x)*\max(x)})}$$

where $\min(x)$ and $\max(x)$ are the minimum and maximum abscissa values, respectively.

The inputs are:

```
# spxfit :: Fit a 1-dimensional model to an image or image region
for determination of spectral index.
imagename      =          # Name of the input image(s)
box            =          '' # Rectangular box in
                                # direction coordinate blc, trc.
                                # Default: entire image ("").
region         =          '' # Region of interest. See
                                # help par.region for possible
                                # specifications. Default:
                                # Do not use a region.
chans          =          '' # Channels to use. Channels
                                # must be contiguous. See "help
                                # par.chans" for
                                # examples. Default: all channels ("").
stokes         =          '' # Stokes planes to
                                # use. Planes must be contiguous. Default:
                                # all stokes ("").
axis           =          -1 # The profile axis. Default:
                                # use the spectral axis if one
                                # exists, axis 0 otherwise (<0).
mask           =          '' # Mask to use. See help
                                # par.mask. Default is none.
minpts         =          1  # Minimum number of unmasked
                                # points necessary to attempt
                                # fit.
multifit       =          True # If true, fit a profile
                                # along the desired axis at each
                                # pixel in the specified
                                # region. If false, average the
                                # non-fit axis pixels and do
                                # a single fit to that average
                                # profile. Default False.
spxsol         =          '' # Name of the spectral index
                                # function coefficient solution
                                # image to write.
spxerr         =          '' # Name of the spectral index
                                # function coefficient error
                                # image to write.
model         =          '' # Name of model
                                # image. Default: do not write the model
```

```

        residual      =      ''      # image ("").
                                # Name of residual
                                # image. Default: do not write the
                                # residual image ("").
spxtype              =      'plp'    # Type of function to
                                # fit. "plp" => power logarithmic
                                # polynomial, "ltp" =>
                                # logarithmic transformed polynomial.
spxest               =      []       # Initial estimates for the
                                # spectral index function
                                # coefficients.
spxfix               =      []       # Fix the corresponding spectral index function
                                # coefficients during the fit. True=>hold fixed.
div                  =      0        # Divisor (numerical value or
                                # quantity) to use in the
                                # logarithmic terms of the
                                # plp or ltp function. 0 =>
                                # calculate a useful value on the fly.
wantreturn           =      True     # Should a record summarizing
                                # the results be returned?
logresults           =      True     # Output results to logger?
logfile              =      ''      # File in which to log
                                # results. Default is not to write a
                                # logfile.
sigma                =      -1      # Standard deviation array or image name(s).
    outsigma         =      ''      # Name of output image used
                                # for standard deviation. Ignored
                                # if sigma is empty.

```

for more than a single input image or cube, all images must have the same dimensions along all axes other than the fit axis. `multifit` will perform a per pixel fit, otherwise there will be a single value over the entire region.

6.21 Image Mask Handling `makemask`

`makemask` facilitates the handling of image masks in CASA. There are two basic mask formats: 1) one or more Boolean masks stored internally in the image file, and 2) images with zero and non-zero image values. `makemask` looks like:

```

# makemask :: Makes and manipulates image masks
mode              =      'list'      # Mask method (list, copy, expand, delete, setdefaultmask)
    inpimage      =      ''         # Name of input image.

```

To distinguish between Boolean internal masks and zero/non-zero images, `makemask` uses the syntax `galaxy.image:mask0` for Boolean masks within an image, in this example the Boolean mask `mask0`

within the image *galaxy.image*. Without the colon separator, the image itself is assumed to be treated like a zero/non-zero mask.

`mode='list'` lists all the internal Boolean masks that are present in an image. The default masks can be set with mode `setdefaultmask` and they can be deleted with the mode `delete`. The default mask is used when an image is displayed in the `viewer` and is used in all analysis tasks.

`mode='copy'` lets a user to copy a Boolean mask from one image to another image, or to write out the Boolean mask as a zero/non-zero image. The latter format is very useful when different masks are combined or manipulated. All the image analysis tools, in particular `immath` are applicable for such zero/non-zero masks as they act like normal images. `makemask` will always attempt to regrid the input mask to the output image.

In addition `mode='copy'` can be used to merge masks and also accepts regions. E.g. to create a mask from a CASA region file (CRTF, see § 6.1.6), the input would look like:

```
# makemask :: Makes and manipulates image masks
mode           =      'copy'           # Mask method (list, copy, expand, delete, setdefaultmask)
inpimage       = 'inputimage.im'      # Name of input image.
inpmask        = 'region.crtf'        # mask(s) to be processed: image masks, T/F internal masks
                                           # (Need to include parent image names), regions (for copy mode)
output         = 'imagemask.im'       # Name of output mask (imagename or imagename:internal_maskname)
overwrite      =      False           # overwrite output if exists?
```

`mode='expand'` furthermore expands a mask in the spectral domain. It regrids first then stretches the edge channels. E.g. a one plane continuum image would be stretched to all planes of a data cube.

6.22 Search for Spectral Line Rest Frequencies (`slsearch`)

The `slsearch` task allows the spectral line enthusiast to find their favorite spectral lines in subset of the Splatalogue spectral line catalog (<http://www.splatalogue.net>) which is distributed with CASA. In addition, one can export custom catalogs from Splatalogue and import them to CASA using the task `splattotable` (Sect. 6.23) or tool method `sl.splattotable()`. One can even import catalogs with lines not in Splatalogue using the same file format.

The inputs to `slsearch` are as follows:

```
# slsearch :: Search a spectral line table.
tablename      =      ''              # Input spectral line table name to
                                           # search. If not specified, use the
                                           # default table in the system.
outfile        =      ''              # Results table name. Blank means do not
                                           # write the table to disk.
freqrange      = [84, 90]             # Frequency range in GHz.
```

```

species      =      ['']      # Species to search for.
reonly       =      False     # List only NRAO recommended
                                     # frequencies.
chemnames    =      ['']      # Chemical names to search for.
qns          =      ['']      # Resolved quantum numbers to search
                                     # for.
rrlinclude   =      True      # Include RRLs in the result set?
rrlonly      =      False     # Include only RRLs in the result set?
  intensity   =      -1       # CDMS/JPL intensity range. -1 -> do not
                                     # use an intensity range.
  smu2        =      -1       # S*mu*mu range in Debye**2. -1 -> do
                                     # not use an S*mu*mu range.
  loga        =      -1       # log(A) (Einstein coefficient) range.
                                     # -1 -> do not use a loga range.
  eu          =      -1       # Upper energy state range in Kelvin. -1
                                     # -> do not use an eu range.
  el          =      -1       # Lower energy state range in Kelvin. -1
                                     # -> do not use an el range.

verbose      =      True      # List result set to logger (and
                                     # optionally logfile)?
  logfile     =      ''       # List result set to this logfile (only
                                     # used if verbose=True).
  append      =      True     # If true, append to logfile if it
                                     # already exists, if false overwrite
                                     # logfile if it exists. Only used if
                                     # verbose=True and logfile not blank.

wantreturn   =      True     # If true, return the spectralline tool
                                     # associated with the result set.

```

The table is provided in the `tablename` parameter but if it is blank (the default), the catalog which is included with CASA will be used. Searches can be made in a parameter space with large dimensionality:

- `freqrange` Frequency range in GHz.
- `species` Species to search for.
- `reonly` List only NRAO recommended frequencies.
- `chemnames` Chemical names to search for.
- `qns` Resolved quantum numbers to search for.
- `intensity` CDMS/JPL intensity range.
- `smu2` $S\mu^2$ range in Debye².
- `loga` $\log(A)$ (Einstein coefficient) range.

- `e1` Lower energy state range in Kelvin.
- `eu` Upper energy state range in Kelvin.
- `rrlinclude` Include RRLs in the result set?
- `rrlonly` Include only RRLs in the result set?

Notation is as found in the *Splatalogue* catalog.

Example:

Search for all lines of the species HOCN and HOCO+ in the 200-300GHz range:

```
ssearch(outfile="myresults.tbl", freqrange = [200,300],
        species=['HOCN', 'HOCO+'])
```

The task can also return a python dictionary if assigned a variable like:

```
myLines = ssearch(outfile="myresults.tbl", freqrange = [200,300],
                  species=['HOCN', 'HOCO+'])
```

6.23 Convert Exported Splatalogue Catalogs to CASA Tables (`splattotable`)

In some cases the internal spectral line catalog may not contain the lines in which one is interested. In that case, one can export a catalog from *Splatalogue* <http://www.splatalogue.net> or even create their own "by hand" (be careful to get the format exactly right though!). CASA's task `splattotable` can then be used to create a CASA table that contains these lines and can be searched:

```
-----> inp(splattotable)
# splattotable :: Convert a downloaded Splatalogue spectral line list to a casa table.
filenames      =      ['']      # Files containing Splatalogue lists.
table          =      ''        # Output table name.
wantreturn     =      True      # Do you want the task to return a spectralline tool attached to
```

A search in *Splatalogue* will return a catalog that can be saved in a file (look for the "Export" section after the results on the search results page). The exported filename(s) should be entered in the `filenames` parameter of `splattotable`. The downloaded files must be in a specific format for this task to succeed. If you use the *Splatalogue* "Export CASA fields" feature, you should have no difficulties.

6.24 Image Import/Export to FITS

These tasks will allow you to write your CASA image to a FITS file that other packages can read, and to import existing FITS files into CASA as an image.

6.24.1 FITS Image Export (exportfits)

To export your images to fits format use the `exportfits` task. The inputs are:

```
# exportfits :: Convert a CASA image to a FITS file
imasename = '' # Name of input CASA image
fitsimage = '' # Name of output image FITS file
velocity = False # Use velocity (rather than frequency) as spectral axis
optical = False # Use the optical (rather than radio) velocity convention
bitpix = -32 # Bits per pixel
minpix = 0 # Minimum pixel value
maxpix = 0 # Maximum pixel value
overwrite = False # Overwrite pre-existing imasename
dropstokes = False # Drop the Stokes axis?
stokeslast = True # Put Stokes axis last in header?
```

The `dropstokes` or `stokeslast` parameter may be needed to make the FITS image compatible with an external application.

For example,

```
exportfits('ngc5921.demo.cleanimg.image', 'ngc5921.demo.cleanimg.image.fits')
```

6.24.2 FITS Image Import (importfits)

You can also use the `importfits` task to import a FITS image into CASA image table format. Note, the CASA viewer can read fits images so you don't need to do this if you just want to look at the image. The inputs for `importfits` are:

```
# importfits :: Convert an image FITS file into a CASA image
fitsimage = '' # Name of input image FITS file
imasename = '' # Name of output CASA image
whichrep = 0 # If fits image has multiple
# coordinate reps, choose one.
whichhdu = 0 # If its file contains
# multiple images, choose one.
zeroblanks = True # Set blanked pixels to zero (not NaN)
overwrite = False # Overwrite pre-existing imasename
defaultaxes = False # Add the default 4D
# coordinate axes where they are missing
defaultaxesvalues = [] # List of values to assign to
# added degenerate axes when
# defaultaxes=True (ra,dec,freq,stokes)
```

For example, we can read the above image back in

```
importfits('ngc5921.demo.cleanimg.image.fits', 'ngc5921.demo.cleanimage')
```

6.25 Using the CASA Toolkit for Image Analysis

Although this cookbook is aimed at general users employing the tasks, we include here a more detailed description of doing image analysis in the CASA toolkit. This is because there are currently only a few tasks geared towards image analysis, as well as due to the breadth of possible manipulations that the toolkit allows that more sophisticated users will appreciate.

To see a list of the `ia` methods available, use the CASA `help` command:

```
CASA <1>: help ia
-----> help(ia)
Help on image object:

class image(_builtin_.object)
| image object
|
| Methods defined here:
|
| __init__(...)
|     x.__init__(...) initializes x; see x.__class__.__doc__ for signature
|
| __str__(...)
|     x.__str__() <==> str(x)
|
| adddegaxes(...)
|     Add degenerate axes of the specified type to the image' :
|         outfile
|         direction = false
|         spectral  = false
|         stokes
|         linear    = false
|         tabular   = false
|         overwrite = false
|         -----
|
| addnoise(...)
...

|
| unlock(...)
|     Release any lock on the image' :
|     -----
|
| -----
| Data and other attributes defined here:
|
| __new__ = <built-in method __new__ of type object at 0x55d0f20>
|     T.__new__(S, ...) -> a new object with type S, a subtype of T
```

Inside the Toolkit:

The image analysis tool (`ia`) is the workhorse here. It appears in the User Reference Manual as the `image` tool. Other relevant tools for analysis and manipulation include `measures` (`me`), `quanta` (`qa`) and `coordsys` (`cs`).

or for a compact listing use <TAB> completion on `ia.`, e.g.

```
CASA <5>: ia.
Display all 105 possibilities? (y or n)
ia.__class__          ia.deconvolvecomponentlist  ia.ispersistent          ia.reorder
ia.__delattr__       ia.deconvolvefrombeam      ia.lock                  ia.replacemaskedpixel
ia.__doc__           ia.done                    ia.makearray            ia.restoringbeam
ia.__getattr__       ia.echo                   ia.makecomplex          ia.rotate
ia.__hash__          ia.fft                    ia.maketestimage       ia.sepconvolve
ia.__init__          ia.findsources            ia.maskhandler         ia.set
ia.__new__           ia.fitallprofiles         ia.maxfit               ia.setboxregion
ia.__reduce__        ia.fitcomponents         ia.miscinfo            ia.setbrightnessunit
ia.__reduce_ex__     ia.fitpolynomial         ia.modify               ia.setcoordsys
ia.__repr__          ia.fitprofile            ia.moments              ia.sethistory
ia.__setattr__       ia.fromarray              ia.name                 ia.setmiscinfo
ia.__str__           ia.fromascii              ia.newimage             ia.setrestoringbeam
ia.adddegaxes        ia.fromfits               ia.newimagefromarray   ia.shape
ia.addnoise          ia.fromimage              ia.newimagefromfile    ia.statistics
ia.boundingBox       ia.fromrecord             ia.newimagefromfits    ia.subimage
ia.brightnessunit    ia.fromshape              ia.newimagefromimage   ia.summary
ia.calc              ia.getchunk               ia.newimagefromshape   ia.toASCII
ia.calcmask          ia.getregion              ia.open                 ia.tofits
ia.close             ia.getslice               ia.outputvariant       ia.topixel
ia.collapse          ia.hanning                ia.pixelvalue          ia.torecord
ia.continuumsub      ia.haslock                 ia.putchunk            ia.toworld
ia.convertflux       ia.histograms            ia.putregion           ia.twopointcorrelatio
ia.convolve          ia.history                 ia.rebin                ia.type
ia.convolve2d        ia.imagecalc              ia.regrid               ia.unlock
ia.coordmeasures     ia.imageconcat            ia.remove
ia.coordsys          ia.insert                  ia.removefile
ia.decompose         ia.isopen                  ia.rename
```

A common use of the `ia` tool is to do region statistics on an image. The `imhead` task has `mode='stats'` to do this quickly over the entire image cube. The tool can do this on specific planes or sub-regions. Here is an example on how to use the `ia` tool to get on-source and off-source statistics:

```
# The variable clnimage points to the clean image name

# Pull the max and rms from the clean image
ia.open(clnimage)
on_statistics=ia.statistics()
thistest_immax=on_statistics['max'][0]
oldtest_immax = 1.07732224464
print ' Clean image ON-SRC max should be ',oldtest_immax
print ' Found : Max in image = ',thistest_immax
diff_immax = abs((oldtest_immax-thistest_immax)/oldtest_immax)
print ' Difference (fractional) = ',diff_immax

print ''
```

```
# Now do stats in the lower right corner of the image
box = ia.setboxregion([0.75,0.00],[1.00,0.25],frac=true)
off_statistics=ia.statistics(region=box)
thistest_imrms=off_statistics['rms'][0]
oldtest_imrms = 0.0010449
print ' Clean image OFF-SRC rms should be ',oldtest_imrms
print ' Found : rms in image = ',thistest_imrms
diff_imrms = abs((oldtest_imrms-thistest_imrms)/oldtest_imrms)
print ' Difference (fractional) = ',diff_imrms

print ''
print ' Final Clean image Dynamic Range = ',thistest_immax/thistest_imrms
print ''
print ' ===== '

ia.close()
```

Note: If you don't close the file with, e.g., `ia.close()` the file will stay in a 'locked' state. Other processes won't be able to access the file until the file is properly closed.

6.26 Examples of CASA Image Analysis

The data reduction tutorials on casaguides.nrao.edu provide walkthroughs for high and low frequency, spectral line and polarization analysis techniques.

Chapter 7

Visualization With The CASA Viewer

This chapter describes using the CASA Viewer to display data. The Viewer can be started as a stand-alone executable or by the `viewer` task inside a CASA shell. It can display both images and Measurement Sets. We are in the process of splitting the task-level access to the Viewer into two tasks: `imview` for images and `msview` for measurement sets. These tasks offer improved scriptability, giving command line access to many of the viewer features.

7.1 Starting the Viewer

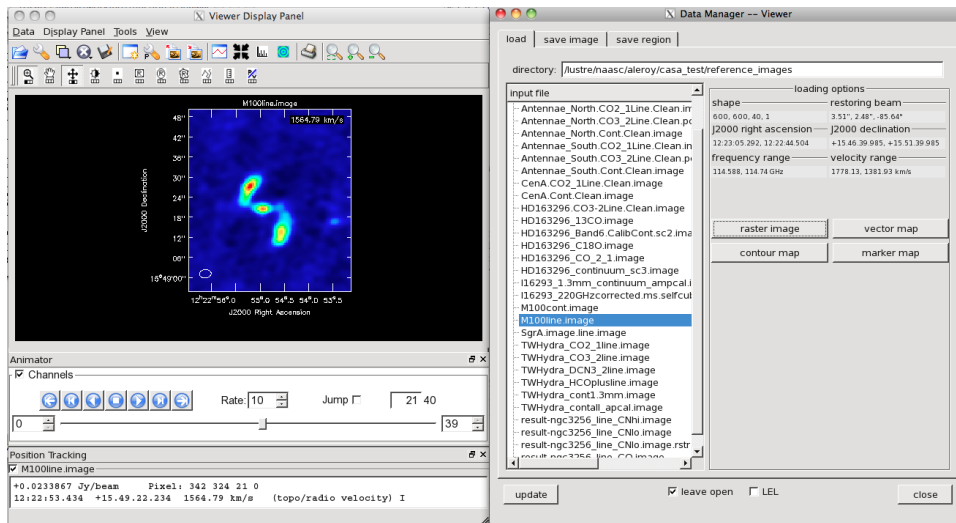


Figure 7.1: The **Viewer Display Panel** (left) and the **Data Manager** (right) panel for a regular image or data cube.

Within the casapy environment, the `viewer` task can be used to start the CASA Viewer, displaying an image or MS. The inputs are:


```
# viewer :: View an image or visibility data set.

infile      =      '' # (Optional) Name of file to visualize.
displaytype = 'raster' # (Optional) Type of visual rendering
                    # (raster, contour, vector or marker).
                    # lel if an lel expression is given
                    # for infile (advanced).
```

Examples of starting the CASA Viewer:

```
CASA <1>: viewer()

CASA <2>: viewer('ngc5921.demo.ms')

CASA <3>: viewer('ngc5921.demo.cleanimg.image')

CASA <4>: viewer('ngc5921.demo.cleanimg.image', 'contour')

CASA <5>: viewer('"ngc5921.demo.cleanimg.image"^2', 'lel')
```

The first command creates an empty **Viewer Display Panel** (§ 7.2) and a **Load Data** window (§ 7.3.1). The second starts the CASA Viewer and loads a Measurement Set. The third example starts the Viewer and opens an image data cube (see Figure 7.1).

Examples four and five make use of the second parameter (`displaytype`). Example four displays the image as a contour map rather than the default raster map. Example five uses 'Lattice (Image) Expression Language' to display the square of the image data.

Note that the Viewer can open FITS files, CASA image files, Measurement Sets, and saved viewer states. The Viewer determines the type of file being opened automatically.

For additional scripting options when opening the Viewer, see the discussion of the `imview` and `msview` tasks at the end of this Chapter (§7.7 and 7.8).

7.1.1 Running the CASA Viewer outside casapy

If you have CASA installed, then the CASA Viewer is available as a stand-alone application called `casaviewer`. From the operating system prompt, the following commands work the same as the `casapy` task commands given in the previous Section:

```
casaviewer &

casaviewer ms_filename &

casaviewer image_filename &

casaviewer image_filename contour &

casaviewer '"image_filename"^2' lel &
```

7.2 The Viewer Display Panel

The CASA Viewer consists of a number of graphical user interface (GUI) windows. The main **Viewer Display Panel** (§ 7.2) is used for both image and Measurement Set viewing. It is shown in the left panels of Figures 7.1 and 7.2 and appears the same whether an image or Measurement Set is being displayed.

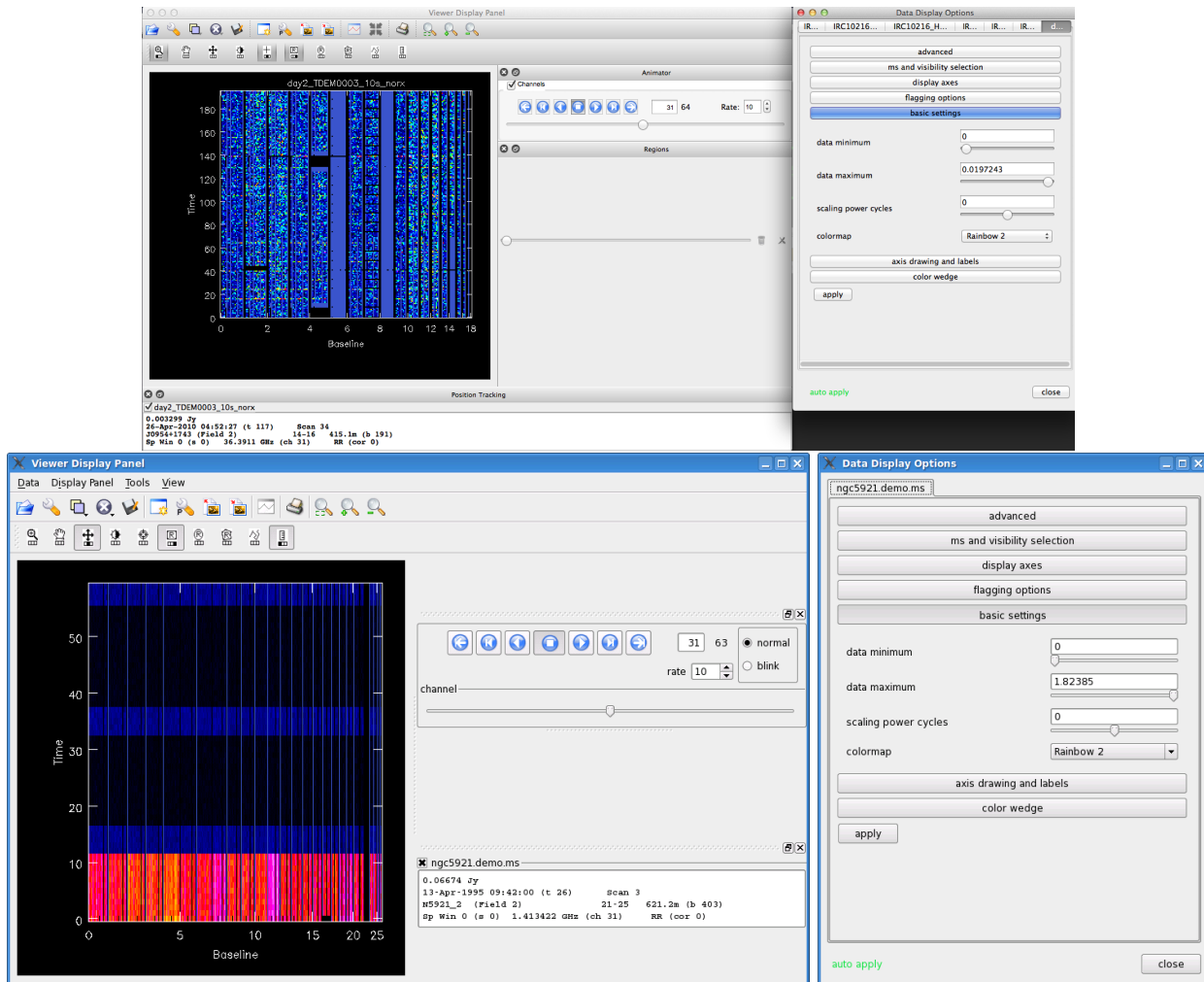


Figure 7.2: The **Viewer Display Panel** (left) and **Data Display Options** (right) panels with a Measurement Set open.

At the top of the Viewer Display Panel are drop down menus:

- **Data**
 - **Open** — open the Data Manager window (§7.3).

- **Register** — select and de-select which of the loaded data file(s) should be displayed. The menu expands to the right showing all loaded data sets. Unchecking an image will cause it not to be displayed, but does not close it.
- **Close** — close (unload) the selected data file. The menu expands to the right showing all loaded data.
- **Adjust Data Display** — open the Data Display Options ('Adjust') window (§7.4.1).
- **Save as...** — save/export data to a file
- **Print** — print the displayed image
- **Save Panel State** — to a 'restore' file (xml format)
- **Restore Panel State** — from a restore file
- **Preferences** — manually edit the viewer configuration
- **Close Panel** — close this Viewer Display Panel. If this is the last display panel open, this will exit the Viewer.
- **Quit Viewer** — close all display panels and exit

- **Display Panel**

- **New Panel** — create a new, empty Viewer Display Panel
- **Panel Options** — open the Viewer Canvas Manager window to edit margins, the number of panels, and the background (§7.4.1.5).
- **Save Panel State** — save the current state of the viewer as a file that can later be reloaded.
- **Restore Panel State** — restore the viewer to a state previously saved as a file.
- **Print** — print displayed image
- **Close Panel** — close this Viewer Display Panel. If this is the last display panel open, this will exit the Viewer.

- **Tools**

- **Spectral Profile** — Open the Spectral Profile Browser window to look at intensity as a function of frequency for part of an image (§7.4.4.2)
- **Collapse Image** — Open the Collapse/Moments window, which allows you to create new images from a data cube by integrating along the spectral axis (§7.4.6)
- **Histogram** — Open the Histogram inspection window, which allows you to graphically examine the distribution of pixel values in a data cube (§7.4.7)
- **Fit** — Open the two-d fitting window, which can be used to fit Gaussians to two dimensional intensity distributions (§7.4.8).
- **Interactive Clean** — Open a window to look at ongoing interactive clean processes.

- **View**

- **Main Toolbar** — show/hide the top row of icons (Figure 7.3, §7.2.1).

- **Mouse Toolbar** — show/hide the second row of mouse-button action selection icons (Figure 7.4, §7.2.2).
- **Animator** — show/hide tapedeck control panel attachment to the main Viewer Display Panel (§7.2.3).
- **Position Tracking** — show/hide the position tracking attachment to the main Viewer Display Panel (§7.2.3).
- **Regions** — show/hide the region manager attachment to the main Viewer Display Panel (§7.2.3).

7.2.1 The Main Toolbar



Figure 7.3: The display panel’s **Main Toolbar** appears directly below the menus and contains ‘shortcut’ buttons for most of the frequently-used menu items.

Below the drop down menus is the **Main Toolbar** (Figure 7.3). This top row of icons offers fast access to these menu items:

- **folder** (`Data:Open` shortcut) — open the Data Manager window (§7.3)
- **wrench** (`Data:Adjust` shortcut) — open the Data Display Options (‘Adjust’) window (§7.4.1).
- **panels** (`Data:Register` shortcut) — select and de-select which of the loaded data file(s) should be displayed. The menu expands to the right showing all loaded data sets. Unchecking an image will cause it not to be displayed, but does not close it.
- **delete** (`Data:Close` shortcut) — close (unload) the selected data file. The menu expands to the right showing all loaded data.
- **save data** (`Data:Save as`) — save the current data to a file.
- **new panel** (`Display Panel:New Panel`) — create a new, empty Viewer Display Panel
- **panel wrench** (`Display Panel:Panel Options`) — open the Viewer Canvas Manager window to edit margins, the number of panels, and the background (§7.4.1.5).
- **save panel** (`Display Panel: Save Panel State`) — save panel state to a ‘restore’ file
- **restore panel** (`Display Panel: Restore Panel State`) — restore panel state from a restore file
- **spectral profile** (`Tools: Spectral Profile`) — Open the Spectral Profile Browser window to look at intensity as a function of frequency for part of an image (§7.4.4.2)

- **collapse/moments**(Tools: Collapse Image) — Open the Collapse/Moments window, which allows you to create new images from a data cube by integrating along the spectral axis (§7.4.6)
- **histogram**(Tools:Histogram) — Open the Histogram inspection window, which allows you to graphically examine the distribution of pixel values in a data cube (§7.4.7)
- **fitting**(Tools:Fit) — Open the two-d fitting window, which can be used to fit Gaussians to two dimensional intensity distributions (§7.4.8).
- **print** (Display Panel:Print) — print the current display
- **magnifier box** — zoom out all the way
- **magnifier plus** — zoom in (by a factor of 2)
- **magnifier minus** — zoom out (by a factor of 2)

7.2.2 The Mouse Toolbar



Figure 7.4: The '**Mouse Tool**' Bar allows you to assign how mouse buttons behave in the image display area. Initially, zooming, color adjustment, and rectangular regions are assigned to the left, middle and right mouse buttons. Click on a tool with a mouse button to assign that tool to that mouse button.

Below the Main Toolbar are eleven **Mouse Tool** buttons (Figure 7.4). These allow you to assign what behavior the three mouse buttons have when clicked in the display area. Clicking a mouse tool icon will [re-]assign the mouse button that was clicked to that tool. Black and white squares beneath the icons show which mouse button is currently assigned to which tool.

The mouse tools available from the toolbar are:

(Note that the 'escape' key can be used to cancel any mouse tool operation that was begun but not completed, and to erase a region, point, or other tool showing in the display area.)

- **Zooming (magnifying glass icon):** To zoom into a selected area, press the Zoom tool's mouse button (the **left** button by default) on one corner of the desired rectangle and drag to the desired opposite corner. Once the button is released, the zoom rectangle can still be moved or resized by dragging. To complete the zoom, double-click inside the selected rectangle. If you instead double-clicking *outside* the rectangle, you will zoom *out*.
- **Panning (hand icon):** Press the tool's mouse button on a point you wish to move, drag it to the position where you want it moved, and release. *Note: The arrow keys, Page Up, Page Down, Home and End keys can also be used to pan through your data any time you are zoomed in. (Click on the main display area first, to be sure the keyboard is 'focused' there).*

- **Stretch-shift colormap fiddling (crossed arrows):** This is usually the handiest color adjustment; it is assigned to the **middle** mouse button by default. Note that you can also adjust the color table quantitatively inside the Data Display Options window (§7.4.1).
- **Brightness-contrast colormap fiddling (light/dark sun):** Another tool to adjust the color stretch.
- **Positioning (plus):** This tool can place a point marker on the display to select a position. It is used to flag Measurement Set data or to select an image position for spectral profiles. Click on the desired position with the tool's mouse button to place the point; once placed you can drag it to other locations. You can also place multiple points on the display (e.g. for different spectral profile positions) – remove them by hovering over and hitting **ESC**. Double-click is not needed for this tool. See § 7.4.3.2 for more detail.
- **Rectangle, Ellipse and Polygon region drawing:** The rectangle region tool is assigned to the **right** mouse button by default. As with the zoom tool, a rectangle region is generated by dragging with the assigned mouse button; the selection is confirmed by double-clicking within the rectangle. An ellipse regions is created by dragging with the assigned mouse button. In addition to the elliptical region, also its surrounding rectangle is shown on the display. The selection is confirmed by double-clicking within the ellipse. Polygon regions are created by clicking the assigned mouse button at the desired vertices, clicking the final location twice to finish. Once created, a polygon can be moved by dragging from inside, or reshaped by dragging the handles at the vertices. See § 7.4.3.2 for the uses of this tool.
- **Polyline drawing:** A polyline can be created by selecting this tool. It is manipulated similarly to the polygon region tool: create segments by clicking at the desired positions and then double-click to finish the line. [Uses for this tool are still to be implemented].
- **Distance tool:** After selecting the distance tool by assigning any mouse button to it, distances on the image can conveniently be measured by dragging the mouse with the assigned button pressed. The tool measures the distances along the world coordinate axes and along the hypotenuse. If the units in both axes are [*deg*], the distances are displayed in [*arcsec*].
- **Position-Velocity Diagram:** Use this mouse tool to drag out a position axis that can be used to generate a position velocity diagram in a new viewer panel from the region manager dock.

7.2.3 The Display Area

The main **Display Area** lies below the toolbars. This area shows the image or Measurement Set currently loaded. Clicking the mouse inside the display area allows region or position selection according to the settings in the mouse toolbar.

The Display Area may have up to three attached panels: the **Animator** panel, the **Position Tracking** panel, and the **Regions** panel. These may be displayed or hidden from the "View" dropdown menu in the main Viewer Display Panel. If one of these is missing from your viewer,

check that it is checked "on" in that menu. The panels can also be turned off by clicking the "X" in the top right corner, in which case you will need to use the View menu to get them back.

By default, the three panels appear attached to the main Viewer Display Panel on the right side of the image. They may be dragged to new positions. Each of the three panels can be attached to the left, top, right, or bottom of the main Viewer Display Panel or they can be entirely undocked and left as free-floating panels.

NOTE: Depending on your window manager, windows without focus, including detached panels and tools like the Spectral Profile Browser may sometimes display odd behavior. As a general rule, giving the window focus by clicking on it will correct the issue. If you seem to "lose" a detached panel (like an Animator Panel), then click in the main window to get it back.

NOTE: With all three panels turned on (and especially with several images loaded), the main display panel can sometimes shrink to very small sizes as the panels grow. Try detaching the panels to get the main display panel back to a useful size.

A restart of the viewer will display all docks in the state of a previous viewer session, given that it was closed normally. In addition, the viewer docking can be changed under "Preferences" in the toolbar (Mac OS under the "CASA Viewer" tab on the toolbar, Linux: "Data"). Fig. 7.5 shows an example. Each item can be changed and the input box will only allow accepted input formats. A complete restart is required to apply the changes.

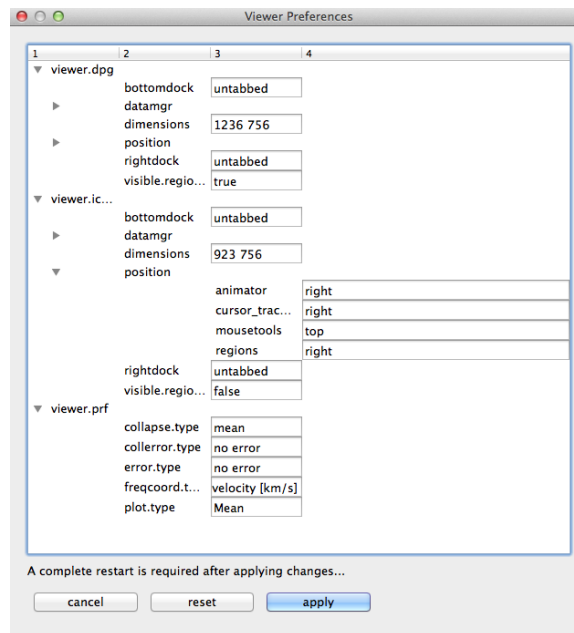


Figure 7.5: "Preferences" dialog to manually change the docking and size of the viewer panel.

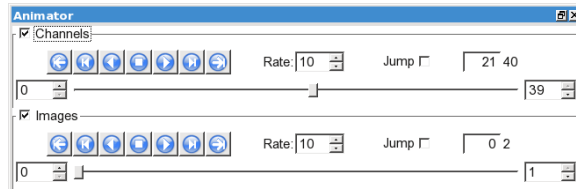


Figure 7.6: The animator panel, which allows one to scroll along the z axis of a data cube (using the Channels tape deck) or cycle among open Images. The panel can be undocked from the main display panel.

7.2.3.1 The Animator Panel

The Animator Panel allows you to scroll through the channels of a data cube and to rotate among loaded images. The main features of the panel are the two “tape decks,” one labeled “Channels” and one Labeled “Images” (note that you will only see the Images tape deck when multiple images are loaded).

The **Channels** tape deck scrolls between planes of an individual image. By default, the channel tape deck scrolls among frequency planes when R.A. and Declination are the displayed axes (in this case, frequency is the “ z axis”). From outside to inside, the buttons cause the display to jump all the way to the beginning/end of the z axis, cause the viewer to step one plane forward or backward along the z axis, or start a movie. The limits on the z axis can be set manually using the windows at the end of the scroll bar. The scroll bar can also be dragged or the user can jump the display to a manually entered plane by entering the plane into the text bock.

When you have multiple images loaded, the **Images** tape deck cycles through which is image is being displayed. In the movie mode, it allows you continuously click between images. Functionally, the image tape deck works similarly to the channels tape deck, with the ability to step, jump, or continuously scroll through images.

NOTE: The check boxes next to the channel and images tabs enable or disable those panels. This doesn’t have much effect when the display has only a single panel, but with multiple panels (i.e., several maps at once in the main window) it changes the nature of the display. If the “Images” box is checked then interleaved maps from different cubes are display. Otherwise a series of maps from a single cube are shown.

7.2.3.2 The Position Tracking Panel

The **Position Tracking** panel (below the images in Fig 7.1) shows the intensity, position (e.g., RA and Dec), Stokes, frequency (or velocity), and pixel location for the point currently under the cursor. A separate box appears for each registered image or Measurement Set and you can see the tracking information for each. Tracking can be ‘frozen’ (and unfrozen again) by hitting the space bar when the viewer’s focus is on the main display area (to be sure that this is case first click on the main display area).

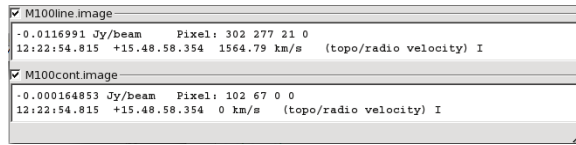


Figure 7.7: The position tracking panel, which gives information about the open data cube at the current location of the cursor. Freeze the position tracking panel using the SPACE bar.

7.2.3.3 The Region Manager Panel

The **Region Manager** panel becomes active when regions are created. It has a large amount of functionality, from display region statistics and histograms to creating position-velocity cuts. We discuss these in §7.4.3. Like the Animator and Position Tracking panel, the Region Manager Panel can be moved relative to the main viewer display panel or entirely undocked.

7.2.4 Saving and Restoring the Display Panel State

You can save the display panel’s current state — meaning the panel settings and the data on display — or load a saved panel state from disk. To save the display panel state, select **Save Panel State** from the **Display Panel** drop-down menu or click the ”Save Display Panel State to File” icon on the main toolbar (an arrow pointing from a picture to a page, see Figure 7.3). It is advisable but not required to retain the file’s `.rstr` (”Restore”) extension.

You can restore the display panel to the saved state by loading the saved state from the Data Manager Panel, by selecting **Restore Panel State** from the **Display Panel** drop down menu, or by clicking the ”Restore Display Panel State” icon (just to the right of the ”Save Display Panel State” icon).

It is possible to restore panel states viewing Measurement Sets or image and panel states that have multiple layers, such as contour plots over raster images. You can also save LEL displays. You can also save or restore the panel state with no data loaded, which is a convenient way to restore preferred initial settings such as overall panel size.

Data Locations: The viewer is fairly forgiving regarding data location when restore a saved panel state. It will find files located:

- in the original location recorded in the restore file
- in the current working directory (where you started the viewer)
- in the restore file’s directory
- in the original location relative to the restore file

This means that you can generally restore a saved panel state if you move that file together with data files. The exception to this rule is that the process is less forgiving if you save the display of an LEL expression. In this case the files must be in the locations specified in the original LEL expression. If a data file is **not** found, restore will attempt to proceed but results may not be ideal.

Manually Editing Saved Display Panel States: The saved "Restore" files are in ascii (xml) format, and manual edits are possible. However, these files are long and complex. Use caution, and back up restore files before editing. If you make a mistake, the viewer may not even recognize the file as a restore file. It is easier and safer to make changes on the display panel and then save the display panel state again.

7.3 The Data Manager Panel — Saving and Loading Data

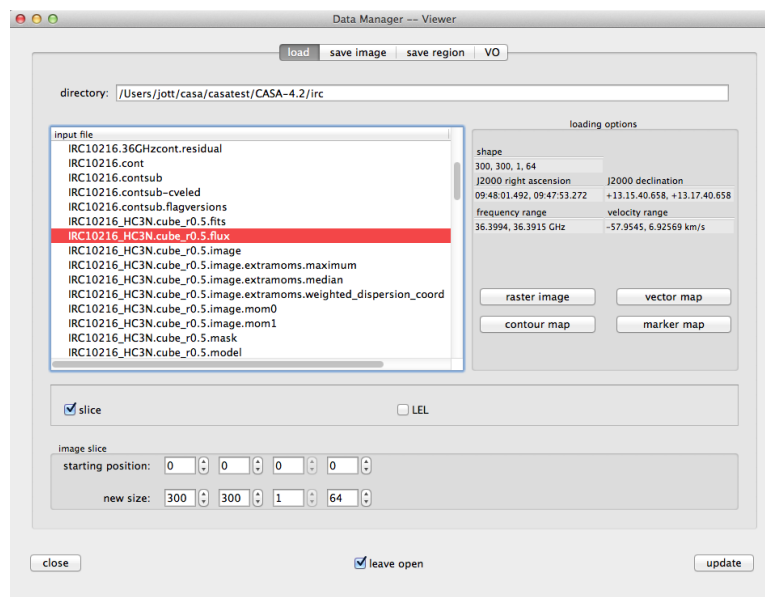


Figure 7.8: The **load** tab of the **Data Manager** panel. This appears if you open the viewer without any **infile** specified, if you use select **Open** from the **Data** drop down menu, or click the Open (Folder) icon. You can access the **save image** or **save region** tabs from this view or by selecting **Save as...** from the **Data** drop down menu. The load tab shows all files in the current directory that can be loaded into the viewer — images, MS, CASA region files, and Display Panel State files.

The **Data Manager Panel** is used to interactively load and save images, Measurement Sets, Display Panel States, and regions. An example of the loading tab in this panel is shown in Figure 7.8. This panel appears automatically if you open the viewer without specifying an input file or it can be accessed through the **Data:Open** menu or Open icon of the **Viewer Display Panel**.

7.3.1 Loading Data

The **load** tab of the **Data Manager Panel** allows you to interactively choose images or Measurement Sets to load into the viewer. The load tab automatically shows you the available images, Measurement Sets, and Display Panel States in the current directory that can be opened by the viewer. When you highlight an image in this view, the tab shows a brief summary of the image: pixel shape, extent of the image on the sky and in frequency/velocity, and restoring beam (if available).

Selecting a file will bring up information about that file in the panel on the right of the Data Manager Panel provide options for how to display the data. Images can be displayed as:

1. raster image
2. contour map
3. vector map
4. marker map

These options are each discussed in § 7.4.

slice: a subselection of a data cube can be loaded, the start and end pixel in each spatial, polarization, and spectral dimension can be selected.

LEL: Instead of only loading an image from disk, you may ask the viewer to evaluate a 'Lattice Expression Language' (LEL) expression (§ 6.1.4). This can be entered in the box provided after you click the "LEL" box. The images used in the LEL expression should have the same coordinates and extents.

Measurement Sets: A Measurement Set can only be displayed as a raster. For measurement sets, the load tab offers options for data selection. This will reduce loading and processing times for visibility flagging.

Regridding Images on Load: Optionally, you may regrid the velocity axis of an image on load to match the current coordinates grid in the Display Panel. In this case, the viewer will interpolate (using the selected interpolation scheme) the cube on disk to share the same velocity gridding as the loaded coordinates. This can be used, e.g., to overlay contour maps of different spectral lines or to make synchronized movies of multiple cubes. Note that the regridding depends on the rest frequency in the image, which is used to calculate the velocities used in regridding.

7.3.2 Registered vs. Open Datasets

When you load data as described above, it is first *opened*, and then *registered* on all existing **Display Panels**.

An *open* dataset has been prepared in memory from disk. All open datasets will have a tab in the **Data Display Options** window, whether currently registered or not.

When a data set is *registered* to a Display Panel its coordinates are aligned to the master coordinate image in the panel and it is ready for drawing. If multiple Display Panels are open then a data set may be registered on one Display Panel and not on another. Only those data sets registered on a particular Display Panel show up in its **Position Tracking** panel.

Why Register More Than One Image? It is useful to have more than one image registered on a panel if you are displaying a contour image over a raster image (§ 7.4.2.1), to 'blink' between images (see **Animator** in § 7.2), or to compare images using the position tracking panel.

Unregistering Images: A data set can be registered or unregistered using the **Image Manager** in the **Data** drop down menu or the **Image Manager** icon (third from left). It will open The Image Manager window and the checkboxes can be used to register or unregister an image.

Closing vs. Unregistering: You can close a data set that is no longer needed using the **Close** option in the **Data** drop-down menu, the "Close" icon (fourth from left), or right mouse button "Close" selection in the Image Manager (§ 7.3.3).

If you close a dataset, you must reload it from disk (or recreate it if it's an LEL expression, regridded image, moment or something similar) to see it again. If you unregister a dataset, it will draw immediately if you re-register it, with its options as you have previously set them. In general, close unneeded datasets but unregister those that you intend to use again.

7.3.3 Image Manager

The **Image Manager** is used to define the master coordinate image, the sequence of images (e.g. for blinking), to register and unregister images, close images, change between raster, contour, vector, and marker displays, and to modify the properties of images. The panel can be invoked from the "Manage Images" tool, the third icon from the left (two overlapping squares).

An example is shown in Fig. 7.9. In this case, four images are loaded into the **viewer**. The sequence of images can be changed by dragging and dropping the images to new positions in the stack. The letter to the left indicates whether the image is a **R**aster, **C**ontour, **V**ector, or **M**arker image. **MC** marks the coordinate master, in this case the second image. The checkboxes are to change the registration statuses. The Coordinate Master image can be defined by a right mouse click, and selection the corresponding option. The right mouse menu button also offers options for quick changes between contour and raster images and to close an image.

The **Options** button will open a drop down box (as shown in Fig. 7.9 for the first image), in which one can again change between image type, change to a different rest frequency (or "Reset" to the value in the image header), or open the "Display Options" panel for that specific image with all the adjustment options explained in (§7.4.1 or §7.4.2).

7.3.4 Saving Data or Regions

The viewer can create new images by carrying out velocity regridding, evaluating an LEL expression, or collapsing a data cube. You can save these images to disk using the Data Manager Panel. Select

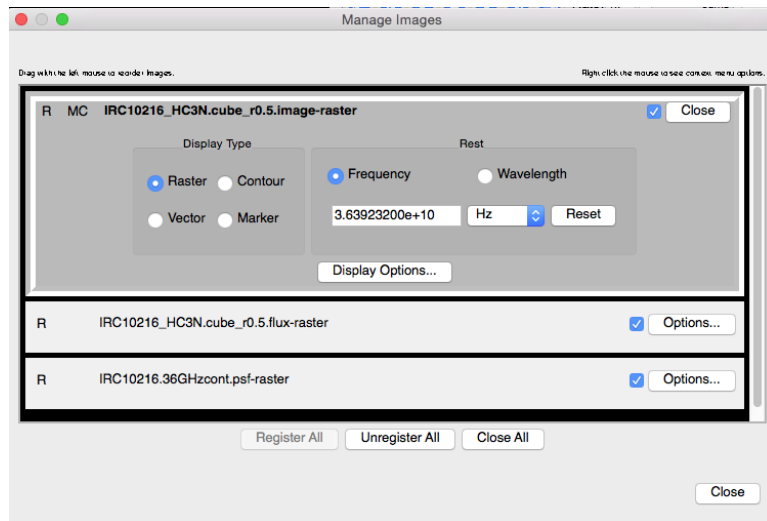
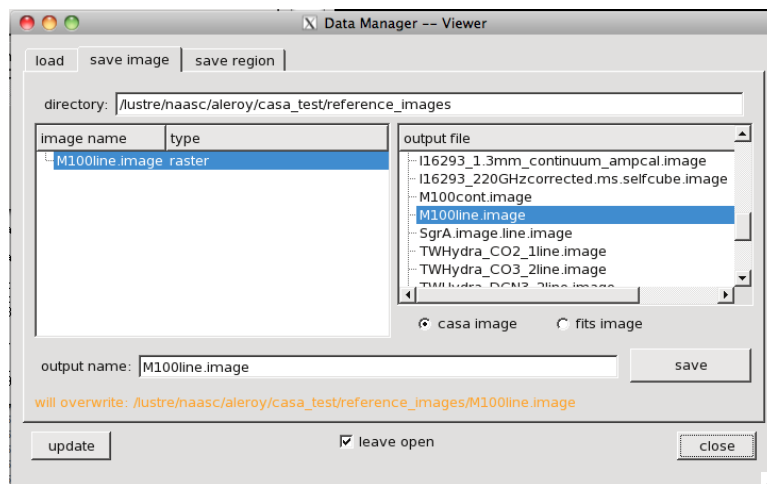


Figure 7.9: The Image Manager.

Figure 7.10: The **Save Data** panel that appears when selecting the 'Save as...' (Figure 7.3).

Save as under the **Data** drop-down menu or click the **Save as** (disk) icon to bring up the Data Manager Panel set to the save tabs. This tab is shown in Figure 7.10.

From the Save Image tab of the Data Manager Panel, you can export images from the viewer to either a CASA image or FITS file on disk. Select the desired file name and click "save." The Data Manager also allows you to save your current regions to a file, either in the CASA or ds9 format. The left part lists all images that can be exported to disk. To save an image to a file, the user can either enter the new filename in the box labeled 'output name:' followed by the save-button

(alternatively the 'Enter'-key), or choose a file name from the right hand side.

7.4 Viewing Images

There are several options for viewing an image. These are seen at the right of the **Load Data - Viewer** panel described in § 7.3.1 and shown in Figure 7.1 after selecting an image. They are:

- **raster image** — a greyscale or color image,
- **contour map** — contours of intensity as a line plot,
- **vector map** — vectors (as in polarization) as a line plot,
- **marker map** — a line plot with symbols to mark positions.

The **raster image** is the default image display, and is what you get if you invoke the Viewer with an image file and no other options. In this case, you will need to use the **Open** menu to bring up the **Load Data** panel to choose a different display.

7.4.1 Viewing a Raster Map

A raster map of an image shows pixel intensities in a two-dimensional cross-section of gridded data with colors selected a colormap according to a scaling that can be specified by the user.

Starting the `casaviewer` with an image as a raster map will look something like the example in Figure 7.1.

Once loaded, the data display can be adjusted by the user through the **Data Display Options** panel, which appears when you choose the **Data:Adjust** menu or use the wrench icon from the **Main Toolbar**.

The **Data Display Options** window is shown in the right panel of Figure 7.1. It consists of a tab for each image or MS loaded, under which are a cascading series of expandable categories. For an image, these are:

- **display axes**
- **hidden axes**
- **basic settings**
- **position tracking**
- **axis labels**
- **axis label properties**

- beam ellipse
- color wedge

The **basic settings** category is expanded by default. To expand a category to show its options, click on it with the left mouse button.

7.4.1.1 Data Display Options — display and hidden axes

In this category the physical axes (i.e. Right Ascension, Declination, Velocity, Stokes) to be displayed can be selected and assigned to the x, y, and z axes of the display. The z axis will be the axis scrolled across by the channel bar in the Animator Panel.

If your image has a fourth axis (typically Stokes), then one of the axes will need to be hidden and not used in viewing. Which axis is hidden can be controlled by a slider within the **hidden axes** drop-down.

7.4.1.2 Data Display Options — basic settings

This roll-up is open by default showing some commonly-used parameters that alter the way the image is displayed. The most frequently used of these change how the intensity value of a pixel maps to a color on the screen. An example of this part of the panel is shown in Figure 7.11.

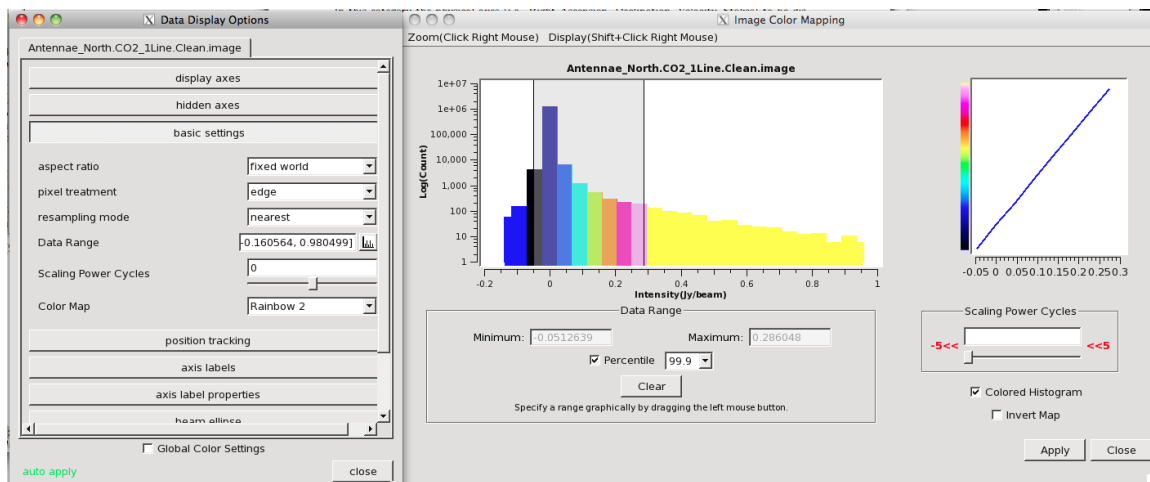


Figure 7.11: The **basic settings** category of the **Data Display Options** panel and the interactive tool for setting the mapping from intensity to color.

The options available are:

- **basic settings: aspect ratio**

This option controls the horizontal-vertical size ratio of data pixels on screen. **fixed world** (the default) means that the aspect ratio of the pixels is set according to the coordinate system of the image (i.e., true to the projected sky). **fixed lattice** means that data pixels will always be square on the screen. Selecting **flexible** allows the map to stretch independently in each direction to fill as much of the display area as possible.

- **basic settings: pixel treatment**

This option controls the precise alignment of the edge of the current 'zoom window' with the data lattice. **edge** (the default) means that whole data pixels are always drawn, even on the edges of the display. For most purposes, **edge** is recommended. **center** means that data pixels on the edge of the display are drawn only from their centers inwards. (Note that a data pixel's center is considered its 'definitive' position, and corresponds to a whole number in 'data pixel' or 'lattice' coordinates).

- **basic settings: resampling mode**

This setting controls how the data are resampled to the resolution of the screen. **nearest** (the default) means that screen pixels are colored according to the intensity of the nearest data point, so that each data pixel is shown in a single color. **bilinear** applies a bilinear interpolation between data pixels to produce smoother looking images when data pixels are large on the screen. **bicubic** applies an even higher-order (and somewhat slower) interpolation.

- **basic settings: data range**

You can use the entry box provided to set the minimum and maximum data values mapped to the available range of colors as a list `[min, max]`. For very high dynamic range images, you will probably want to enter a **max** less than the data maximum in order to see detail in lower brightness-level pixels.

NOTE: By default you edit the scaling of a single image at once and can click between the tabs at the top of the Data Display Options window to manipulate different windows. By checking the Global Color Settings box at the bottom of this window, you will manipulate the scaling for all images at once. This can be very useful, for example, when attempting detailed comparison multiple reductions of the same data.

- **basic settings: scaling power cycles**

This option allows logarithmic scaling of data values to colormap cells, which can be very helpful in the case of very high dynamic range.

The color for a data value is determined as follows:

1. The value is clipped to lie within the data range `[min, max]` specified above.
2. This clipped value is mapped to a new value depending on the selected scaling power cycles in the following way:
 - If the scaling power cycles is set to 0 (the default), the program considers a linear range from `[min, max]` and scales this directly onto the set of available colors.

- For negative scaling values, the data value (after clipping on $[\min, \max]$) is scaled linearly to lie between 0 and 10^p (where p is the value chosen) and then program takes the logarithm of this values, yielding a value in the range 1 to p . That value is scaled linearly to the set of available colors. Thus the data is treated as if it had p decades of range, with an equal number of colors assigned to each decade.
 - For positive scaling values, the data value (after clipping on $[\min, \max]$) is scaled linearly to lie between 0 and p (where p is the value chosen) and 10 is raised to this power, yielding a value in the range 1 to 10^p . That value is scaled linearly to the set of available colors.
3. The color corresponding to a number in final range is determined by the selected colormap and its 'fiddling' (shift/slope) and brightness/contrast settings (see **Mouse Toolbar**, above). Adding a **color wedge** to your image can help clarify the effect of the various color controls.

See Figure 7.12 for sample curves.

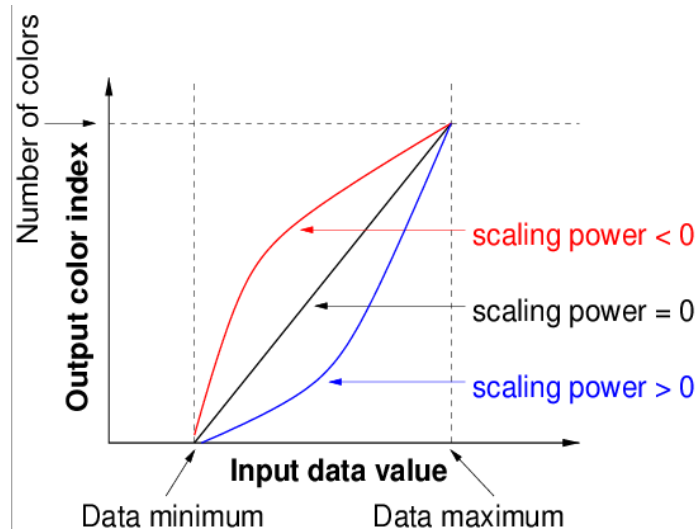


Figure 7.12: Example curves for scaling power cycles.

In practice, you will often manipulate the data range bringing the max down in high dynamic range images, raising the minimum to the near the noise level when using non-zero scaling cycles. It is also common to use negative power cycles when considering high dynamic range images — this lets you bring out the faint features around the bright peaks.

- **basic settings:** colormap

You can select from a variety of colormaps here. **Hot Metal**, **Rainbow** and **Greyscale** colormaps are the ones most commonly used.

7.4.1.3 Graphical Specification of the Intensity Scale

A histogram icon next to the data range in the Data Display opens the Image Color Mapping window, which allows visualization and graphical manipulation of the mapping of intensity to color. The window at the left shows a histogram of the data with a gray range showing the data range. You can use this window to select the data range graphically (with the mouse), manually (by typing into the empty windows), or as a percentile of the data. On the right, you can select the scaling power cycles and see a visualization of the transfer function mapping intensity (x -axis) to color (y -axis).

The functionality here follows the other histogram tools, with the Display tab used to change the histogram plotting parameters. It will often be useful to use a logarithmic scaling of the y -axis and filled histograms when manipulating the color table.

7.4.1.4 Data Display Options — other settings

Many of the other settings on the Data Options panel for raster images are self-explanatory such as those which affect `beam ellipse` drawing (only available if your image provides beam data), or the form of the `axis labeling` and `position tracking` information. You can also give your image a `color wedge`, a key to the current mapping from data values to colors.

7.4.1.5 Viewer Canvas Manager — Panels, Margins, and Backgrounds

The display area can also be manipulated from the **Viewer Canvas Manager** window. Use the wrench icon with a 'P' (or the 'Display Panel' menu) to show this window, which allows you to manipulate the infrastructure of the main display panel. You can set:

- Margins - specify the spacing for the left, right, top, and bottom margins
- Number of panels - specify the number of panels in x and y and the spacing between those panels.
- Background Color - white or black (more choices to come)

Figure 7.13 illustrates a multi-panel display along with the Viewer Canvas Manager settings which created it.

7.4.2 Viewing a Contour Map

Viewing a contour image is similar to viewing a raster map. A contour map shows lines of equal data value for the selected plane of gridded data (Figure 7.14). Contour maps are particularly useful for overlaying on raster images so that two different measurements of the same part of the sky can be shown simultaneously (§ 7.4.2.1).

Several `basic settings` options control the contour levels used:

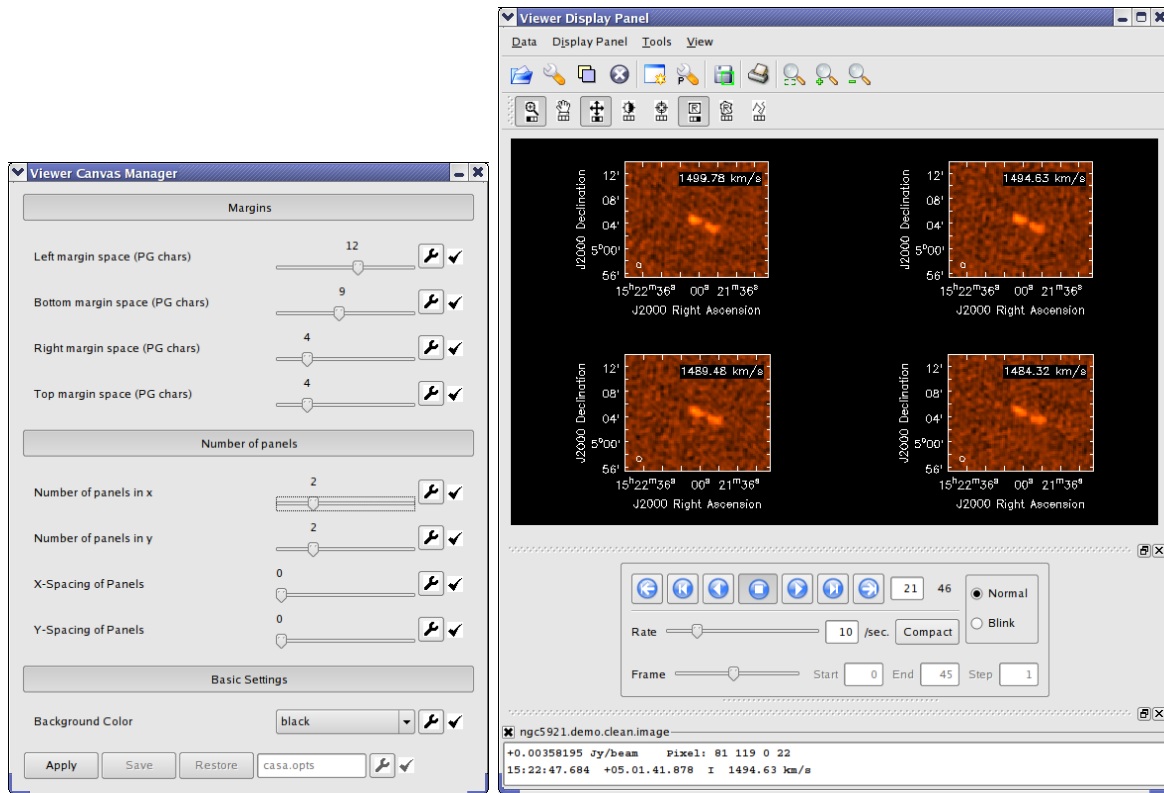


Figure 7.13: A multi-panel display set up through the **Viewer Canvas Manager**.

- The contours themselves are specified by a list in the box **Relative Contour Levels**. These are defined relative to the two other parameters:
- The **Base Contour Level** sets the zero level for the relative contour list corresponds to in units of intensity in the image.
- The **Unit Contour Level** sets what 1 in the relative contour list corresponds to in units of intensity in the image.

Additionally, you have the option to manipulate the thickness and color of the image and to have either positive or negative contours appear dashed.

For example, the following settings:

```
Relative Contour Levels = [0.2, 0.4, 0.6, 0.8]
Base Contour Level = 0.0
Unit Contour Level = <image max>
```

would map the maximum of the image to 1 in the relative contour levels and the base contour level to zero. So the contours will show 20%, 40%, 60%, and 80% of the peak.

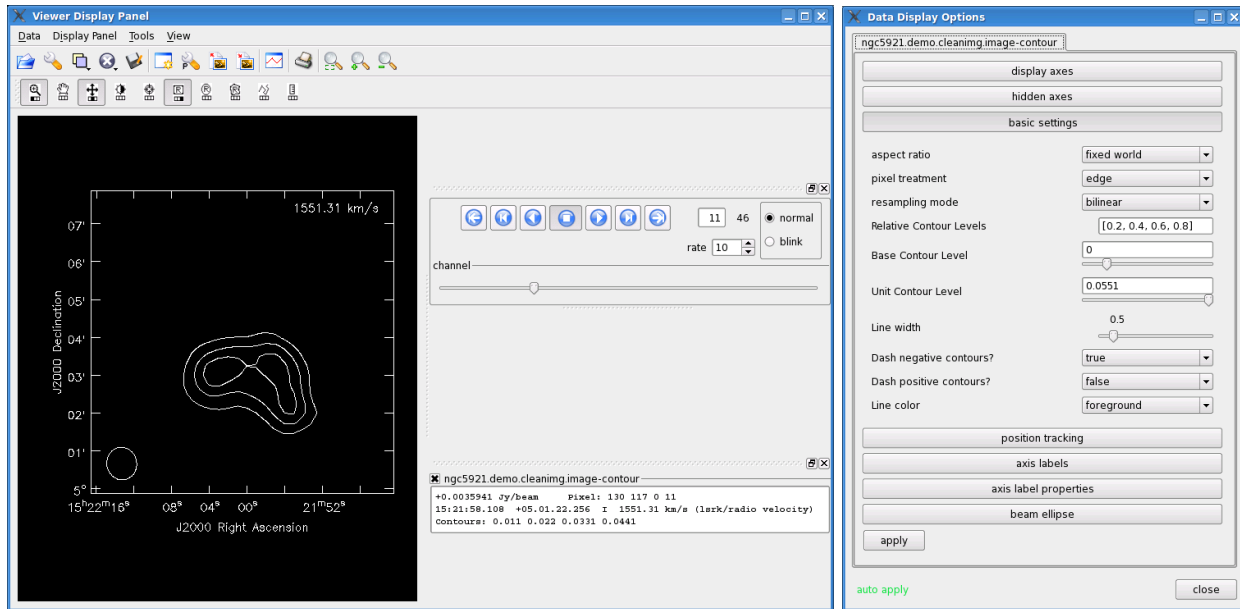


Figure 7.14: The **Viewer Display Panel** (left) and **Data Display Options** panel (right) after choosing contour map from the **Load Data** panel. The image shown is for channel 11 of the NGC5921 cube, selected using the **Animator** tape deck, and zoomed in using the tool bar icon. Note the different options in the open **basic settings** category of the **Data Display Options** panel (as compared to **raster image** in Figure 7.1).

Another approach is to set the unit contour to 1, so that the contours are given in intensity units (usually Jy/beam). So this setup:

```
Relative Contour Levels = [0.010, 0.0.020, 0.040, 0.080, 0.160, 0.320]
Base Contour Level = 0.0
Unit Contour Level = 1.0
```

would create contours starting at 10 mJy/beam and doubling every contour.

Another useful approach is to set contours in units of the rms noise level of the image, which can be worked out from a signal free region. Then a setup like this:

```
Relative Contour Levels = [-3,3,5,10,15,20]
Base Contour Level = 0.0
Unit Contour Level = <image rms>
```

Would indicate significance in the image. The first two contours show emission at ± 3 -sigma and so on.

You can get the image rms using the `imstat` task (§ 6.9) or using the Viewer statistics tool on a region of the image (§ 7.4.3.3).

Not all images are of intensity, for example a moment-1 image (§ 6.7) has units of velocity. In this case, absolute contours (like the last two examples) will work fine, but by default the Viewer will set fractional contours but referred to the min and max of the image:

```
Relative Contour Levels = [0.2, 0.4, 0.6, 0.8]
Base Contour Level = <image min>
Unit Contour Level = <image max>
```

Here we have contours spaced evenly from min to max, and this is what you get by default if you load a non-intensity image (like the moment-1 image). See Figure 7.15 for an example of this.

7.4.2.1 Overlay Contours on a Raster Map

Contours of either a second data set or the same data set can be used for comparison or to enhance visualization of the data. The Data Options Panel will have multiple tabs (switch between them at the top of the window) that allow you to adjust each overlay individually.

NOTE: `axis labeling` is controlled by the *first-registered* image overlay that has labeling turned on (whether raster or contour), so make label adjustments within that tab.

To add a Contour overlay, open the **Load Data** panel (Use the **Data** menu or click on the folder icon), select the data set and click on `contour map`. See Figure 7.15 for an example using NGC5921.

7.4.3 Regions and the Region Manager

CASA regions are following the CASA 'crtf' standard as described in § D. CASA regions can be used in all applications, including `clean` and image analysis tasks (§ 6).

NOTE: A leading 'ann' (short for annotation) to a region definition indicates that it is for visual overlay purposes only.

NOTE: Whereas the region format is supported by all the data processing tasks, some aspects of the `viewer` implementation are still limited to rectangles, ellipses, and some markers. Full support for all region types is progressing with each CASA release.

Once one or more regions are created, the Region Manager Panel becomes active (see Figure 7.16). Like the Position Tracking and Animator Panels, this can be docked or detached from the main viewer display. It contains several tabs that can be used to adjust, analyze, and save or load regions.

NOTE: Moving the mouse into a region will bring it into focus for the Spectral Profile or Histogram tools.

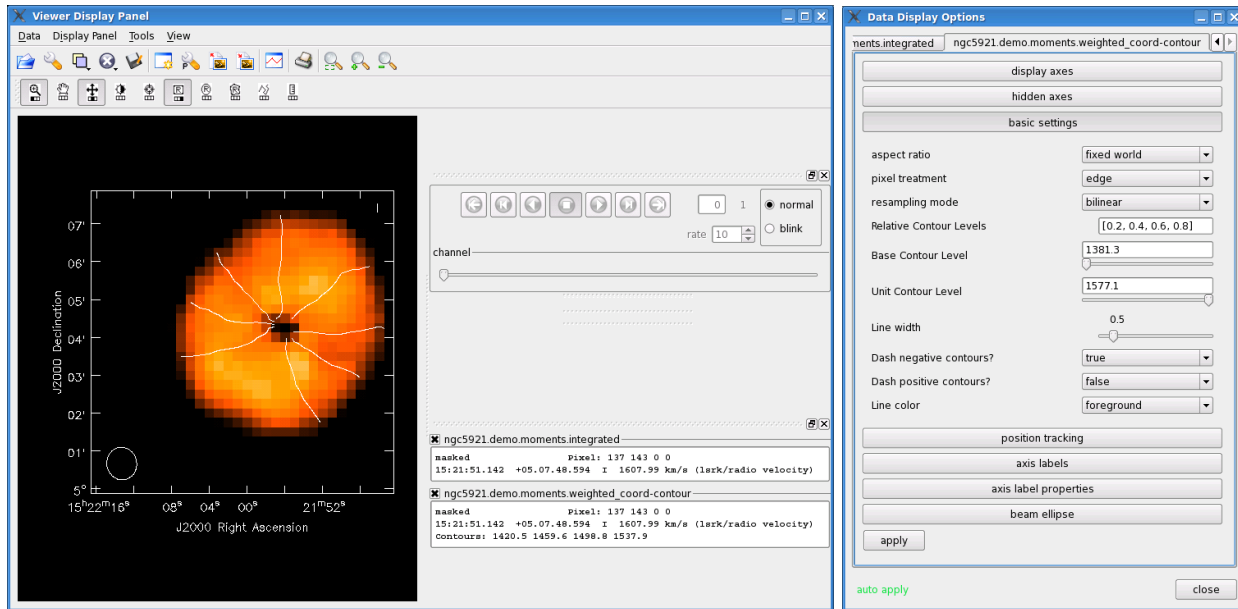


Figure 7.15: The **Viewer Display Panel** (left) and **Data Display Options** panel (right) after overlaying a **Contour Map** of velocity on a **Raster Image** of intensity. The image shown is for the moments of the NGC5921 cube, zoomed in using the tool bar icon. The tab for the contour plot is open in the **Data Display Options** panel.

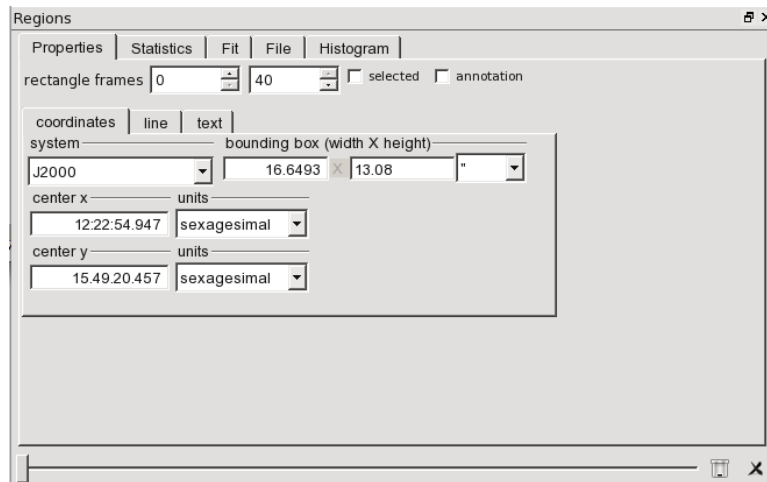


Figure 7.16: The **Region Manager Panel**, which becomes active once at least one region is created. Cycle through available regions using the slider bar at the bottom and use the various tabs to adjust, analyze, load, and save regions.

7.4.3.1 Region Creation, Selection, and Deletion

Within the display area, you can draw regions or select positions using the mouse. Regions can be created with the buttons marked as 'R' in the mouse tool bar (§ 7.2, § 7.4.3.2). The viewer currently supports creation of rectangles, ellipses, polygons, and the point. As usual, a mouse button can be assigned to each button as indicated by the small black square in each button (marking the left, middle, or right mouse button § 7.2, § 7.4.3.2). An example is shown in Fig. 7.17.

Regions can be selected by SHIFT+click, de-selected by pressing SHIFT+click again. The bottom of the Region Manager Panel features a slider to switch between regions in the image. Regions can be removed by hovering over and pressing ESC or by pressing the buttons to the right side of the slider where the first button deletes all regions and the far right button deletes the region that is currently displayed in the panel.

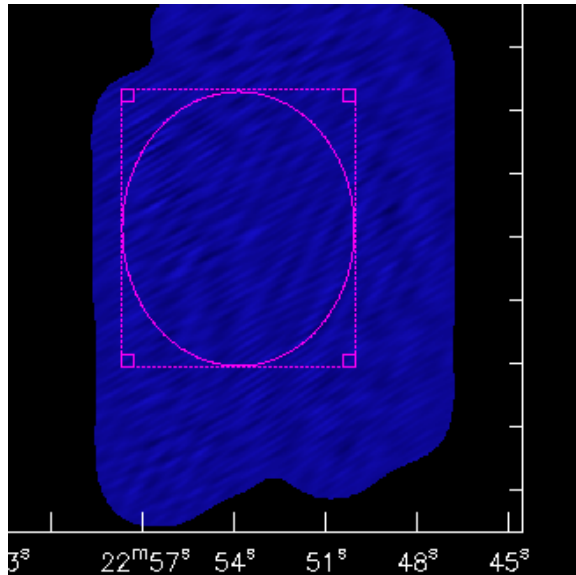


Figure 7.17: Selecting an image region (done with SHIFT+click). The region can be resized by dragging the handles or deleted by hitting ESCAPE.

Once regions are selected, they will feature little, skeletal squares in the corners of their boundary boxes. Selected regions can be moved by dragging with the mouse button and manually resize by grabbing the corners as handles. If more than one region is selected, all selected regions move together.

The **Rectangle Region** drawing tool currently enables the full functionality of the various Region Manager tabs (see below) as well as:

- Region statistics reporting for images via double clicking (also shown in the stats tab of the Region Manager),

- Defining a region to be averaged for the spectral profile tool (accessed via the `Tools:Spectral Profile` drop down menu or "Open the Spectrum Profiler" icon),
- Flagging of Measurement Sets. Note that the `Rectangle Region` tool's mouse button must also be double-clicked to confirm an MS flagging edit.
- Selecting Clean regions interactively (§ 5.3.5)

The `Polygon Region` and `Ellipse Region` drawing have the same uses, except that polygon region flagging of a Measurement Set is not supported.

7.4.3.2 Region Positioning

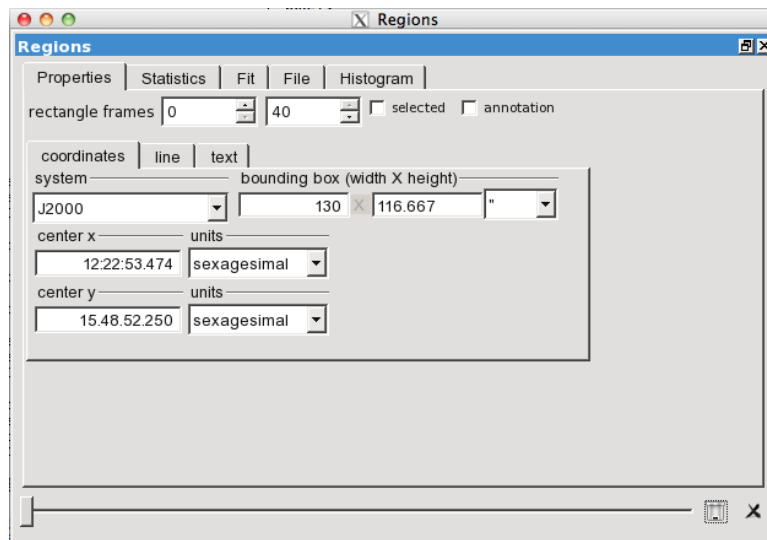


Figure 7.18: The positioning tab in the Region Manager. Use it to manually adjust the location, width, and display style of the selected region.

With at least one region drawn, the region manager becomes active. Using the **Properties** tab, one can manually adjust the position, annotation, and display style of the region. The **frames** boxes set which planes of the image cube the region persists through (regions can have a depth associated with them and will only appear in the frames listed in this range). One can manually adjust the width and height and the center of the box in the chosen units. The **'selection'** check box is an alternative way to the `SHIFT+click` to select a region. The **'annotation'** checkbox will place the **'ann'** string in front of the region ascii output – annotation regions are not be used for processing in, e.g. data analysis tasks. In the **line** and **text** tabs, one can set the style with which the region is displayed, the associated text, and the position and style of that text.

NOTE: Updating the position of a region will update the spectral profile shown if the Spectral Profile tool is open and the histogram if the Histogram tool is open. The views are linked. Dragging a region or adjusting it manually with the Position tab is a good way to explore an image.

7.4.3.3 Region Statistics

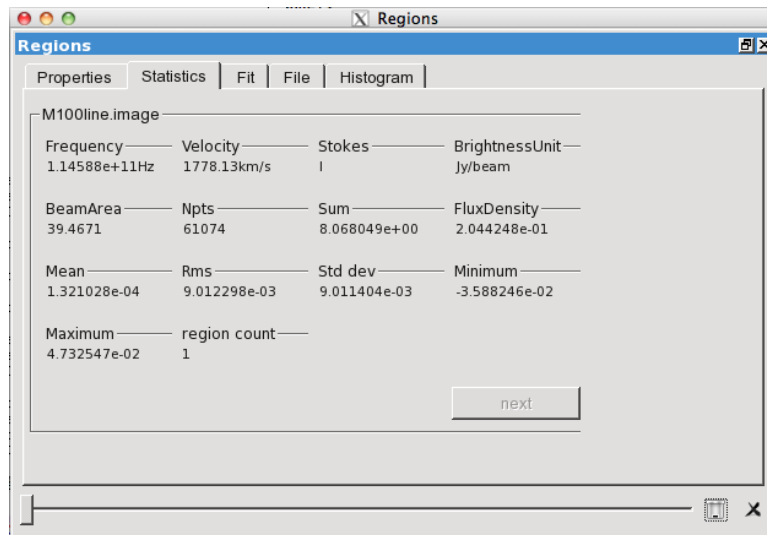


Figure 7.19: The statistics tab in the Region Manager.

One of the most useful features of defining a region is the ability to extract statistics characterizing the intensity distribution inside the region. You can see these in the **Statistics** tab of the of the Region Manager Panel (see Figure 7.19). This displays statistics for the current region in the current plane of the current image. When more than a single region is drawn, you can select them one by one and the region panel will update the statistics to reflect the currently selected region. All values are updated on the fly when the region is dragged across the image.

A similar functionality can be achieved by double clicking inside of a region. This will send statistics information for this region in all registered images to the terminal, looking something like this:

```
(IRC10216.36GHzcont.image) image
      Stokes      Velocity      Frame      Doppler      Frequency
      I -2.99447e+11km/s      LSRK      RADIO      3.63499e+10
BrightnessUnit      BeamArea      Npts      Sum      Flux
      Jy/beam      36.2521      27547      1.087686e-01      3.000336e-03
      Mean      Rms      Std dev      Minimum      Maximum
      3.948473e-06      3.723835e-04      3.723693e-04      -1.045624e-03      9.968892e-03
```

Listed Parameters are *Stokes*, and the displayed channel *Velocity* with the associated *Frame*, *Doppler* and *Frequency* value. *Sum*, *Mean*, *Rms*, *Std Deviation*, *Minimum*, and *Maximum* value refer to those in the selected region and has the units as specified in *BrightnessUnit*. *Npts* is the number of pixels in the region, and *BeamArea* the beam size in pixels. *FluxDensity* is in Jy if the image is in Jy/beam.

This is an easy way to copy and paste the statistical data to a program outside of CASA for further use.

Taking the RMS of the signal-free portion of an image or cube is a good way to estimate the noise. Contrasting this number with the maximum of the image gives an estimate of the dynamic range of the image. The FluxDensity measurement gives a way to use the viewer to do very basic photometry.

7.4.3.4 Saving and Loading Regions

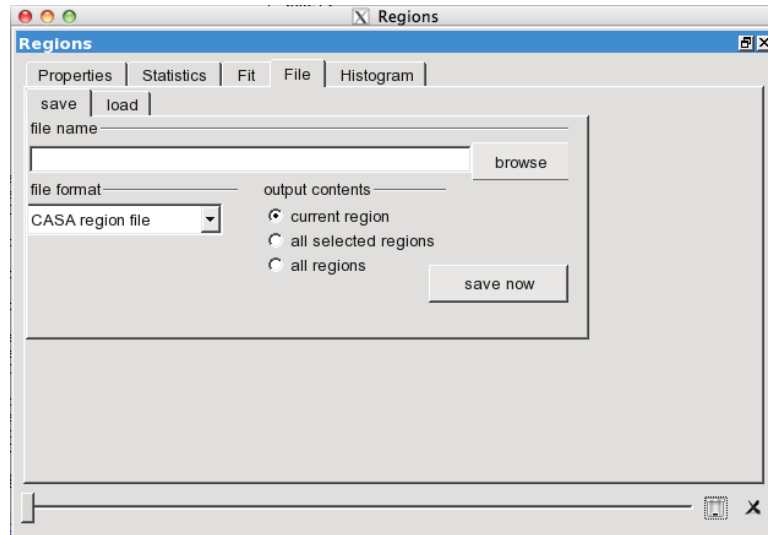


Figure 7.20: The save/load tab in the Region Manager.

The **File** tab in the Region Manager allows one to save or load selected regions, either individually or en masse. You can choose between CASA and ds9 region format. The default is a CASA region file (saved with a *.crtf suffix, see § D). The DS9 format does not offer the full flexibility and cannot capture stokes and spectral axes. DS9 regions will only be usable as annotations in the viewer, they cannot be used for data processing in other CASA tasks. When saving regions, one can choose to save only the current region, all regions that were selected with SHIFT+click, or all regions that are visible on the screen.

NOTE: The load functionality for this tab will only become available once at least one region exists. To load a region when no regions exist, use the Data Manager window (§ 7.3).

7.4.3.5 The Region Fit

NOTE: This functionality is still under development. Its robustness and functionality will be improved in future version of CASA.

The Viewer can attempt to fit a two dimensional Gaussian to the emission distribution inside the currently selected region. To attempt the fit, go to the **Fit** tab of the Region Manager and click the

gaussfit button in the bottom left of the panel. You can choose whether or not to fit a sky level (e.g., to account for a finite background, either astronomical, sky, or instrumental). After fitting the distribution, the **Fit** panel shows the results of the fit, the center, major and minor axis, and position angle of the Gaussian fit in pixels (I) and in world coordinates (W, RA and Dec). The detailed results of the fit will also appear in the terminal window, including a flag showing whether the fit converged.

7.4.3.6 The Region Histogram

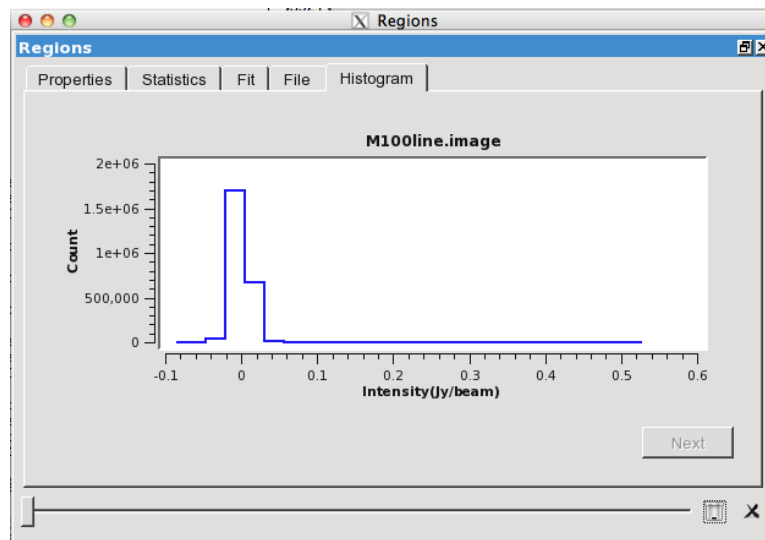


Figure 7.21: The histogram tab in the Region Manager. Right click to zoom. Hit **SHIFT + Right Click** to adjust the details of the histogram display.

The Viewer will automatically derive a histogram of the pixel values inside the selected region. This can be viewed using the **Histogram** tab of the of the Region Manager Panel. This is a pared down version of the full Histogram Tool. You can manipulate the details of the histogram plot by clicking:

1. Use the **Right Click** to zoom - either to the full range, a selected percentile, or a range that you have graphically selected by dragging the mouse (may still be under development).
2. Hit **SHIFT + Right Click** to open the histogram options. This lets you toggle between a logarithmic and linear y-axis, choose between a line, outline, or filled histogram, and adjust the number of bins.

The histogram will update as you change the plane of the cube or shift between region.

7.4.4 The Spectral Profile Tool

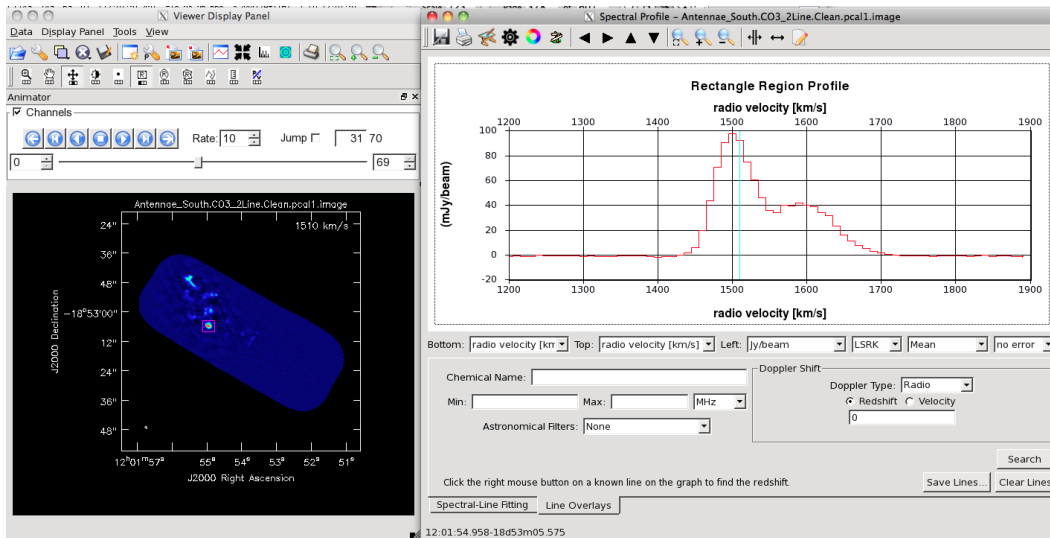


Figure 7.22: The **Spectral Profile** panel (right) that appears when pressing the button **Open the Spectral Profiler** in the **Main Toolbar** and then use the tools to select a region in the image, such as the rectangular region on the left panel. The Spectral Profile tool shows the spectrum of the most recent region highlighted and updates to track movements of the region if moved by dragging with the mouse.

The **Spectral Profile Tool** allows you examine the intensity as a function frequency or velocity. To start a new Spectral Profile window, click the **Spectral Profile** option from the **Tools** drop-down menu or click the "Spectral Profile" (red line graph) icon from the Main Toolbar (see Fig. 7.3). A new Spectral Profile window will appear.

NOTE: *Make Sure That You Use the Radio Version!* This section describes the "Radio" version of the profiler. To be sure that you have the radio version of the tool selected (this may not be the default), click on the preferences icon (the gear fourth from the left) and make sure that the "Optical" option is not checked.

The Spectral Profile Tool consists of a toolbar (§7.4.4.1), a main display area (§7.4.4.2), and two associated tabs: Spectral-Line Fitting (§7.4.4.3) and Line Overlays (§7.4.4.4).

Interaction With the Main Display Panel: For the Spectral Profile tool to work, a region or point must be specified in the main Viewer Display window. Use the mouse tools to specify a point, rectangle, ellipse, or polygon region. Alternatively, load a region file. The Spectral Profile tool will show a spectrum extracted from the region most recently highlight by the mouse in the main Viewer Display Panel. The method of extraction can be specified by the user using a drop down menu below the spectrum in the Spectral Profile window; the method of extraction is mean by default).

The Spectral Profile tool can also feed back to the Main Display Panel. By holding CTRL and right

clicking in the spectrum, you will cause the Main Display Panel to jump to display the frequency channel corresponding to the spectral (x) coordinate of the region clicked in the Spectral Profile tool. Holding CTRL and dragging out a spectral range while holding the right mouse button will queue a movie scrolling through images across that spectral range. You can achieve the same effect with the two-ended-arrow icon towards the right of the toolbar in the Spectral Profile window.

In both tabs, it will be useful to do select regions of frequency or velocity. You can do this with the parallel lines-and-arrow icon (see below) or by holding shift, left clicking, and dragging out the range of interest. A shaded gray region should appear indicating your selection.

7.4.4.1 Spectral Profile Toolbar



Figure 7.23: The toolbar for the Spectral Profile tool allows the user to save the spectrum, print or save the tool as an image, edit preferences (general, tool, legend), spectral smoothing, pan or zoom around the spectrum, select a range of interest, jump to a channel, or add a label.

Figure 7.23 shows the toolbar from the top portion of the **Spectral Profile** window. From left to right, the icons allow the user to:

- (disk) export the current profile to a FITS or ASCII file
- (printer) print the main window to a hard copy
- (writing desk) save the panel as an image (PNG, JPG, PDF, etc.)
- (gear) set plot preferences
- (color wheel) set color preferences for the plot
- (signpost) set legend preferences
- (triangle) set the spectral smoothing method and kernel width
- (arrows) pan the spectrum in the indicated direction **NOTE:** The arrow keys also allow one to pan using the keyboard.
- (magnifying glass) zoom to the default zoom, in, and out **NOTE:** the +/- keys allow one to zoom with the keyboard
- (parallel lines+arrows) drag out a range of interest in the spectrum, for use with fitting or line overlays.

- (double-ended arrow) jump to a channel in the main viewer (single click) or define a range over which to play a movie in the viewer (with a drag). **NOTE:** You can also jump to a channel with **CTRL+Right Click** and queue a movie by holding **CTRL** and dragging out a range while holding the right mouse button.
- notepad and pencil Add or edit a label on the plot. Click this icon to enter a mode where you can drag out a box to create a new annotation box or drag the corners of an existing one to resize it. You can edit the contents, color, and font of an existing annotation by right clicking on it and selecting "Edit Annotation" in the main Spectral Profile window.

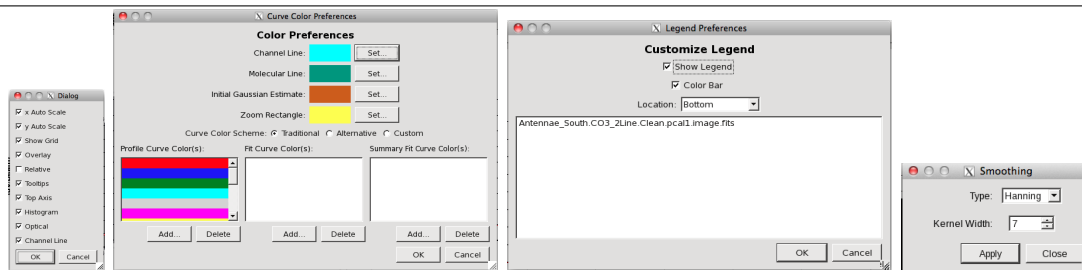


Figure 7.24: Preferences options in the Spectral Profile Tool. From the toolbar, one can access dialogs to set overall viewer preferences, colors for plotting, how the plot legend is displayed, and spectral smoothing method and kernel width.

Figure 7.24 shows the setting dialogs accessible from the toolbar. This **Preferences** dialog opened by the gear icon allows the user to:

- Toggle automatic scaling the x- and y-ranges of the plot.
- Toggle the coordinate grid overlay in the background of the plot.
- Toggle whether registered images other than the current one appear as overlays on the plot.
- Toggle whether these profiles are plotted relative to the main profile (*in development*).
- Toggle the display of tooltips (*in development*).
- Toggle the plotting of a top axis.
- Toggle between a histogram and simple line style for the plot.
- Toggle between the radio and optical versions of the Spectral Profile tool *Note: We discuss only the radio version here; this mainly impacts the Spectral Line Fitting and Collapse/Moments functionality..*
- Toggle the overplotting of a line showing the channel currently being displayed in the main Display Panel.

The **Color Curve Preferences** dialog opened by the color wheel icon allows the user to:

- Select the color of the line marking the current channel shown in the main Display Panel.
- Select the color used to overlay molecular lines from Splatalogue.
- Select the color to plot the initial Gaussian estimate used in spectral line fitting.
- Select the color used for the zoom rectangle.
- Set a queue of colors used to plot the various data sets registered in the Display Panel.
- Set a queue of colors to plot the set of Gaussian fits.
- Set a queue of colors to plot the synthesized curve.

Two sets of preset colors, "Traditional" or "Alternative", are available and the user can define their own custom color palette.

The legend options opened by the signpost icon allow the user to toggle the plotting of a legend defining the curves shown in the main Spectral Profile window. Using a drop-down dialog, the legend can be placed in the top left corner of the plot, to the right of the plot, or below the plot. Toggling the color bar causes the color of the curve to be indicated either via a short bar or using the color of the text itself. Double click the names of the files or curves to edit the text shown for that curve by hand. A legend is only available if more than a single spectrum has been loaded.

The spectral smoothing option has two methods, *Boxcar* and *Hanning* with the selection of odd numbers for the smoothing kernel width in channels.

7.4.4.2 Main Spectral Profile Window

The main window shows the spectrum extracted from active region of the image in the main Display Panel. The spectra from the same region in any other registered images are also plotted if overlays are enabled. Menus along the bottom of the image allow the user to select how the spectrum is displayed. From left to right:

- The units for the bottom spectral axis.
- The units for the top spectral axis. **NOTE:** dual axes are enabled only if a single image is registered and the top axis option is enabled. In general, dual axes are not well-defined for mixed data sets. The exception is that open data cubes with matched frequency/spectral axes will allow dual axes.
- The units for the left intensity or flux axis *Note: fraction of peak allows easy comparison of data with disparate intensity scales..*
- The velocity reference frame used if a velocity axis is chosen for the top or bottom axis.

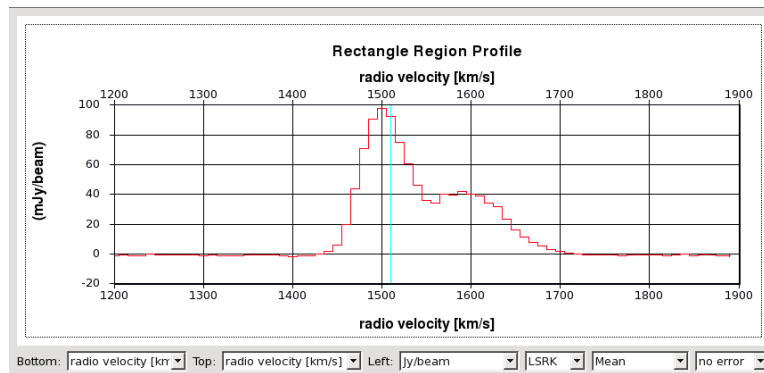


Figure 7.25: The main panel for the **Spectral Profile** tool. Buttons along the bottom row allow the axes to be set. Arrow keys pan and dragging out an area with the mouse zooms. Holding CTRL and right clicking in the spectrum will jump the main Viewer Display panel to display that frequency channel.

- The method used to extract spectrum from the region — a mean over all pixels in the region, a median, sum, or a sum converting units to get a flux density over the region (Jy).
- Toggle the calculation and overplotting of error bars calculated from scatter in the data (`rmse` refers to root mean square error).

In addition to these drop-down menus, the main Spectral Profile window allows the user to do the following using keyboard and mouse inputs:

- *jump the main Display Panel window to a specified channel (CTRL+Right click)*: hold CTRL and right click in the spectrum. A marker will appear and the main Viewer Display Panel will jump to display that channel.
- *animate the main Display Panel in a movie across a frequency range (CTRL+Right click+drag)*: hold CTRL, Right click, and drag. The main Viewer Display panel will respond by showing a movie scrolling across the selected spectral channels.
- *zoom the Spectral Profile (+/-, mouse drag)*: Use the +/- keys to zoom in the same way as the toolbar buttons. Alternatively, press and dragging the left mouse button. A yellow box is drawn onto the panel. After releasing the mouse button, the plot will zoom to the selected range.
- *pan the Spectral Profile (arrows)*: Use the arrow keys to pan the plot.
- *select a spectral range for analysis*: hold shift, left click, and drag. A gray area will be swept out in the display. This method can be used to select a range for spectral line fitting or collapsing a data cube (in the Collapse/Moments window).

NOTE: If the mouse input to the **Spectral Profile** browser becomes confused hit the **ESC** key several times and it will reset.

7.4.4.3 Spectral-Line Fitting

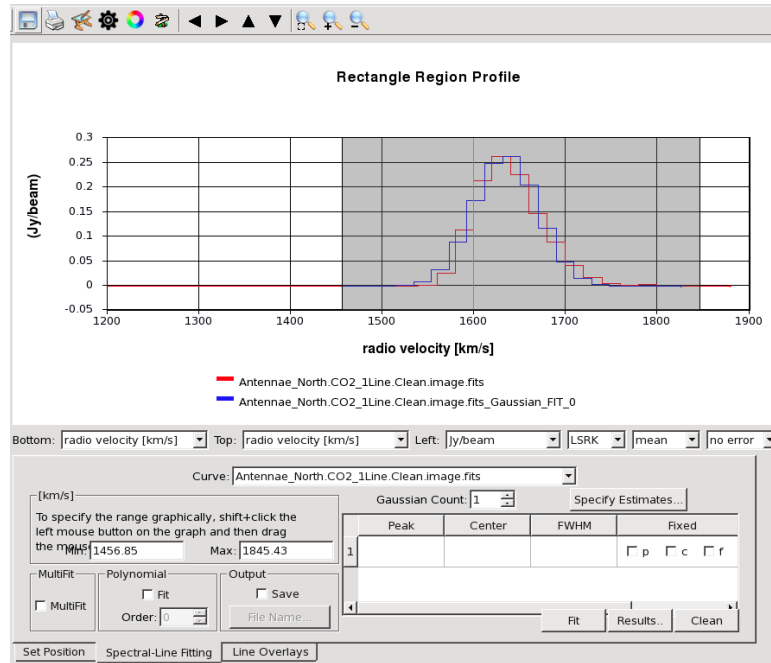


Figure 7.26: The Spectral Line Fitting tab in the Spectral Profile Tool. The user can fit a combination of a polynomial and multiple Gaussian components, specifying the range to be fit (gray region) manually or with a shift+click+drag. Initial estimates for each component may be entered by hand or specified via an initial estimates GUI. The results are output to a dialog and text file with the fit overlotted (here in blue) on the spectrum (with the possibility to save it to disk).

NOTE: Interactive spectral line fitting is still under development.

The **Spectral-Line Fitting** tab, shown in Figures 7.26 and 7.27, allows the user to interactively fit a combination of Gaussian and polynomial profiles to the data shown in the Spectral Line Profile tool. The tool includes a number of options, many of which remain under development:

- A drag-down menu at the top of the panel allows the user to pick which data set to fit.
- The spectral range to fit can be specified by either holding shift+left click+dragging out a region in the main spectral profile window or by typing it manually into the box at the top left of the fitting panel.
- Optionally multiple fits can be carried out once, fitting each spectrum in the region in turn. To enable this, check the “MultiFit” box. (*Under development.*)

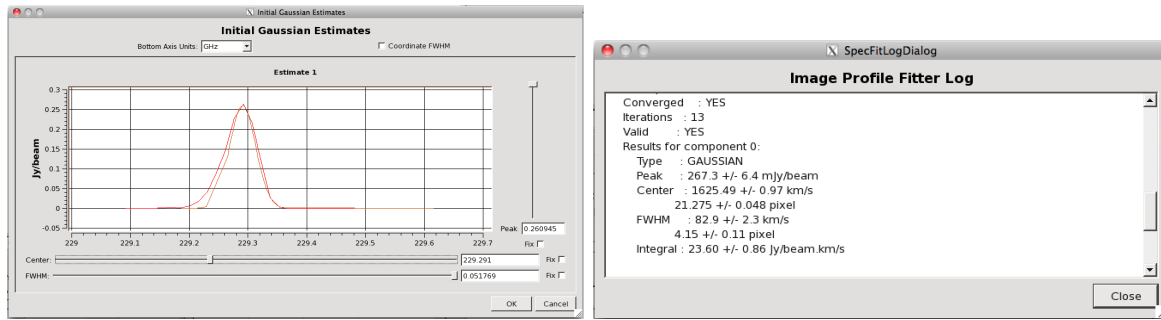


Figure 7.27: The left panel shows the graphical specification of initial estimates for Gaussian fitting. Slider bars specify the center, FWHM, and peak intensity for the initial estimate. The right panel shows the verbose output of the fitting.

- Optionally a polynomial of the specified order may be fit. To do so, check the "Polynomial" fit check box and then specify the desired order. (*Under Development.*)
- The results may be saved to a text file. This text file should be specified *before* the fit is carried out. Click "Save" and then use the dialog to specify the file name. *Note that the fit curve itself becomes a normal spectral profile data set and can be saved to disk using the toolbar (disk icon) after the fit.*
- One or more Gaussians can be fit (*Results are presently most stable for one Gaussian.*). Specify the number of Gaussians and then enter initial estimates for the peak, center, and FWHM in the table below. Any of these values can be fixed for any of the Gaussians being fit. Initial estimates can also be manually specified by clicking **Specify Estimates**. This brings up an additional GUI window (Figure 7.27), where slides can be used to specify initial estimates for each Gaussian to be fit.
- For plotting purposes, one may wish to oversample the fit (i.e., plot a smooth Gaussian), you can do so by increasing the Fit Samples/Channel to a high number to finely sample the fit when plotting.

NOTE: Currently the tool works well for specifying a single Gaussian. Fitting multiple components can become unstable and polynomial and multiple line-of-sight fitting are still under development. This is an area of active development and future releases will offer improved capabilities.

7.4.4.4 Line Overlays

CASA ships with a local version of the Splatalogue spectral line database (www.splatalogue.net) and this can be used to identify and overplot spectral transitions. This feature, shown in Figure 7.28, allows the user to search Splatalogue over the range of interest.

To overlay spectral lines:

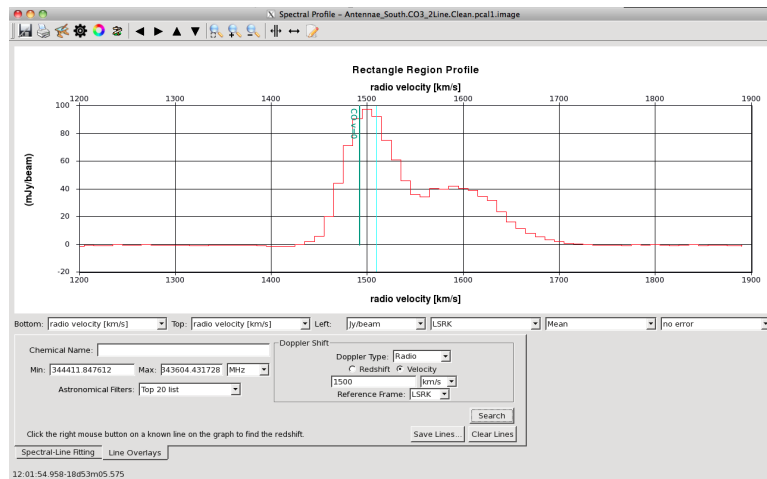


Figure 7.28: Line Overlays in the Spectral Profile Tool. The **Line Overlay** tab, shown at the bottom, allows users to query the CASA copy of the Splatalogue spectral line database. Enter the redshift of your source (right panel), select an Astronomical Filter from the drop down menu, and use shift+click+drag to select a frequency range (or do so manually). The "Search" button will bring up the dialog seen at the left top part of the image, which can in turn be used to graph the candidate lines in the main Spectral Profile window (here CO $v=0$).

1. Select the **Line Overlays** tab in the Spectral Profiles tab.
2. If you know it, enter the redshift or velocity of your source in the "Doppler Shift" panel. Otherwise, the lines will be overlaid assuming a redshift of 0.
3. Specify a minimum and maximum frequency range to search, either by typing a range or by holding shift and left click and dragging out a range in the spectrum (you will see a gray box appear). If you don't specify a range, the tool will search over the frequency range of spectrum.
4. Optionally, you may select an astronomical filter from the list (for example, commonly used extragalactic lines or lines often found in hot cores, see Splatalogue for more information). This is usually a good idea because it pares the potentially very large list of candidate lines to a smaller set of reasonable candidates.
5. Click "Search" and the Spectral Profile will search Splatalogue for a list of Spectral lines that fit that Astronomical Filter in that frequency range for that redshift. A dialog will pop up showing the list of candidate lines.
6. Highlight one or more of these transitions and click "Graph Selected Lines." A set of vertical markers will appear in the main Spectral Profile window at the appropriate (redshifted) frequencies for the line.

We emphasize that this feature remains under active development. Look for improved performance and an expanded feature set in the next release.

NOTE: You will want to click "Clear Lines" between searches, especially if you update the redshift.

7.4.5 The Brightness Profile Tool

The "line" tool can be used to display 1-dimensional brightness profiles of images. The viewer accepts even more than one line segments such as shown in Fig.7.29. The "region" dock will then display a preview of the slice in the "Slice Cut" tab and the full "1-D Slice Tool" can be launched from there. This panel allows one to select the interpolation method to connect the pixels, and a number count for the sampled pixels in between markers. "Automatic" will show all pixels. The x-axis of the display can be either the distance along the slice or the X and Y coordinate projections (e.g. along RA and DEC). All segments are also listed at the bottom with their start and end coordinates, the distance and the position angles of each slice segment. The color tool can be used to give each segment a separate color.

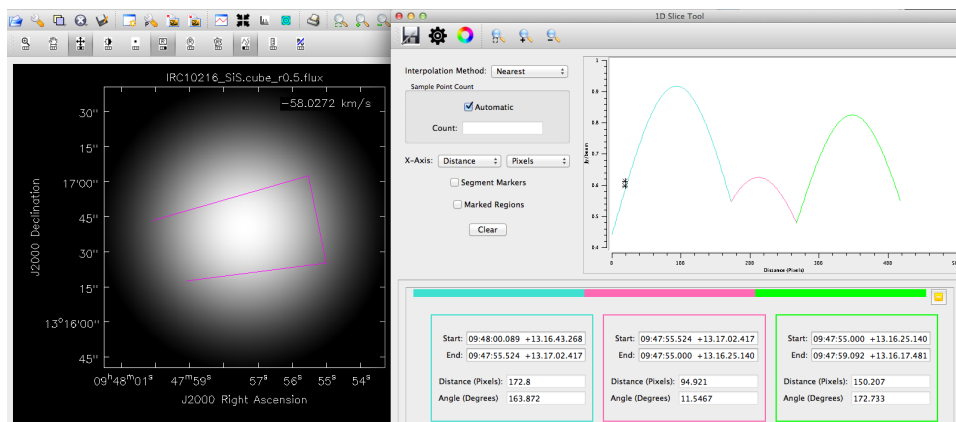


Figure 7.29: 1-dimensional slice of an image. The 1D slicer tool shows the brightness distribution along line segments.

7.4.6 The Collapse/Moments Tool

The CASA Viewer can collapse a data cube into an image, for instance allowing to look at the emission integrated along the z axis or the mean velocity of emission along the line of sight. You can access this functionality via the Collapse/Moments tool (accessed via the Tools menu or the four arrow icons), shown in Figure 7.30.

The tool uses the same format as the Spectral Profile tool and will show the integrated spectrum of whatever region or point is currently selected in the main Display Panel. To create a moment map:

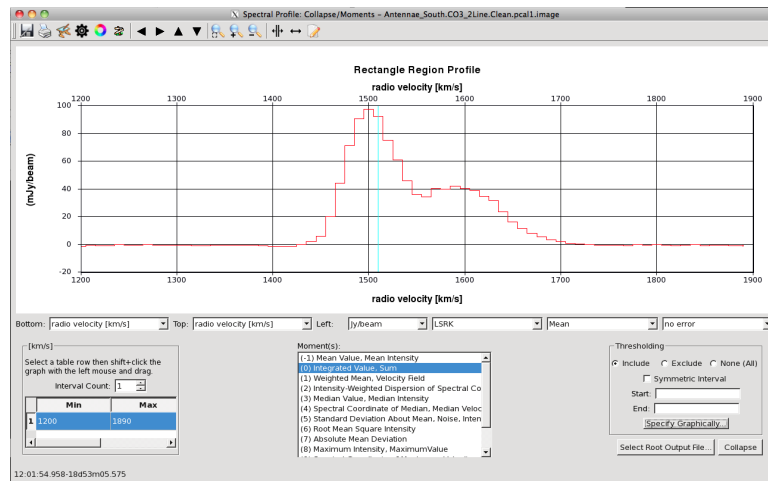


Figure 7.30: The **Collapse/Moments** tool, accessed from the Main Toolbar or the **Tools** drop down menu. The mean spectrum from the region in the Main Display Panel appears in the top part of the tool. After selecting a range, a moment to calculate, and optionally data to exclude click collapse to calculate a new image.

1. Select a range over which to integrate either manually using the left part of the window, by adding an interval and typing in the values into the box or by holding **SHIFT + Left Click** and dragging out the range of interest.
2. Pick the set of algorithms that you will use to collapse the image along the z axis. Clicking toggles each moment method, and the collapse will create a new image for each selected moment. For details on the individual collapse method, see the `immoments` task for more details on each moment.
3. The moment may optionally include or exclude pixels within a certain range (for example, you might include only values with signal-to-noise three or greater when calculating the velocity dispersion). You can enter the values to include or exclude manually in the **Thresholding** window on the right or you can open a histogram tool to specify this range graphically by clicking **Specify Graphically** (before this can work, you must click "Include" or "Exclude").
4. The results of the collapse be saved to a file, which consists of a string specifying the specific moment tacked onto a root file name that you can specify using **Select Root Output File**.
5. When you are satisfied with you chosen options, press **Collapse**.

NOTE: Even if you don't save the results of the collapse to a file, you can still save the map later using the **Save as...** entry in the Data pull down menu from the main Viewer Display Panel.

NOTE: This area remains under active development and may still exhibit some stability issues in CASA 4.1.

7.4.7 The Histogram Tool

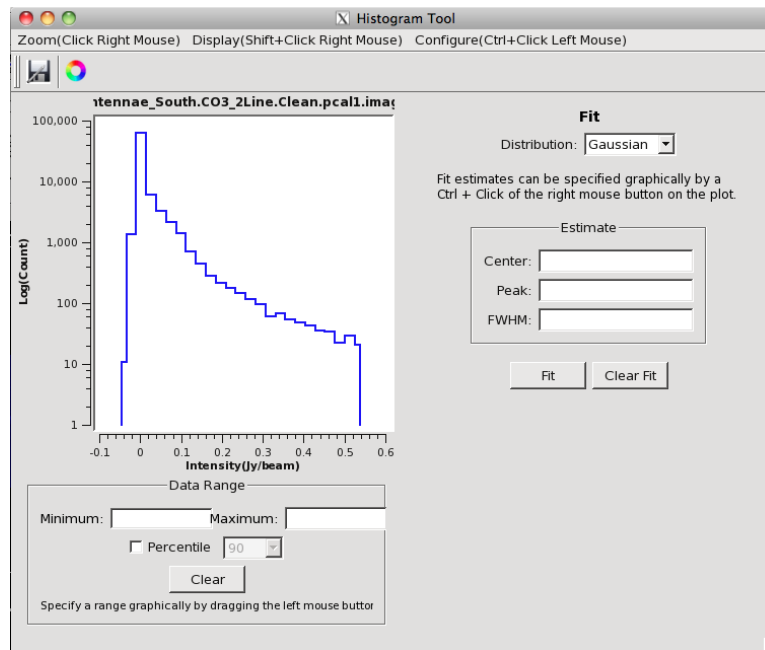


Figure 7.31: The **Histogram** tool, accessed from the Main Toolbar or the **Tools** drop down menu. Details of the display and included pixels can be manipulated via the menus along the top of the window. The right hand panel allows one to attempt to fit a distribution to the histogram.

CASA can calculate and visualize a histogram of pixel values inside a region of interest. To examine this histogram, select **Histogram** from the **Tools** drop-down menu or the Histogram icon (looks like a comb). This opens the full histogram tool; more limited versions are accessible from the Region Manager Panel, the graphical color table manipulation tool, and the Collapse/Moments tool.

The resulting Histogram Tool should look something like Figure 7.31. The menus along the top (or the corresponding mouse clicks) allow one to:

- Zoom to the full range, a selected percentile, or a graphical range.
- Change the display of the histogram to show a log axis, display as either a line plot, an outline, or a filled histogram. Change the number of bins in the histogram, or clear the plot (to start over).
- Configure what data are fed into the histogram. You can use this menu to tell the histogram to track the channel currently selected in the main Viewer Display Panel (click the "Track Channel" box) or to integrate across some range of channels (defaulting to the whole image). You can also switch the 2-D footprint used between the whole Image, the Selected Region, and All Regions.

- Save (via the disk icon) an image of the histogram to a graphical file on disk.

The Histogram Tool also allows you to fit the distribution using either a Gaussian or a Poisson distribution, for example to estimate the noise in the image (a Gaussian will be a good choice to describe the noise in most radio data cubes). You can specify initial estimates or let the program generate initial guesses. The fit is then overplotted on the histogram (colors can be adjusted by clicking the color wheel icon in the toolbar) and details of the fit are printed to the text window below the fit button.

7.4.8 The Two-D Fitting Tool

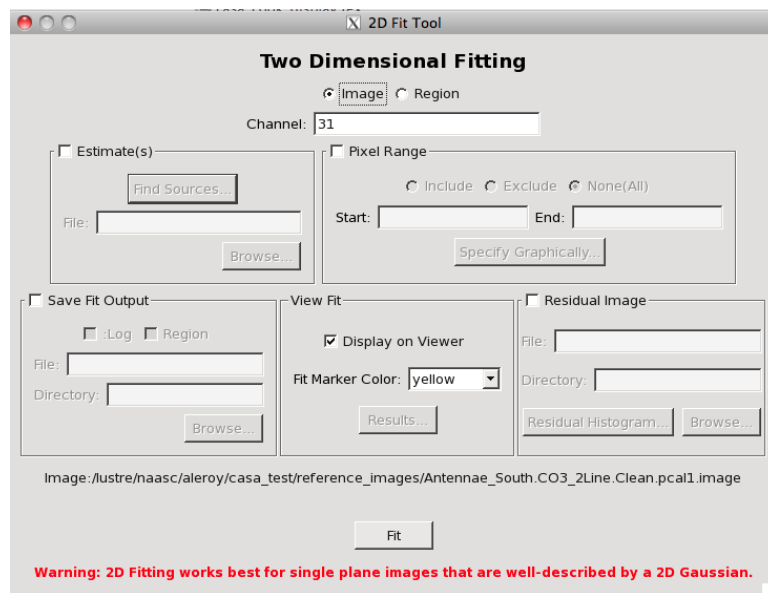


Figure 7.32: The interface to the two dimensional fitting tool (`Tools:Fit...` or the blue circles icon). The interface allows you to specify and automatically generate (**Find Sources**) initial estimates, to specify the range of pixel values to be included in the fit, and to specify the output (log file, residual image, and visualization). Click **Fit** to start the fit.

NOTE: This functionality is still under very active development. Not all features are functional at this point.

CASA can fit two dimensional Gaussians to an intensity distribution, and the Two-Dimensional Fitting tool in the Viewer exposes this functionality interactively. This tool, accessed by the blue circles icon or the `Tools:Fit...` menu item, has an interface like that shown in Figure 7.32. The interface exposes several options:

1. You can select whether to fit only the selected region or the whole image plane and specify

which channel of the cube you want to operate on. **NOTE:** The two dimensional fitter only operates on a single channel at a time.

2. **Initial Estimates:** The box in the top left corner allows the user to specify initial estimates by feeding in a file. The easiest way to make an appropriate file is to edit an existing one. Even easier, you can use the **Find Sources** button to automatically generate a temporary file of initial estimates. **NOTE:** This functionality is still under development. When it is working, you click on **Find Sources**
3. **Pixel Range:** You can choose to only include a certain range of pixel intensity values in the fit. For example, you might choose to only fit Gaussians to pixels above a few times the measured noise level. You can use the **Specify Graphically** button to bring up an interactive histogram of the region (a reduced functionality version of the full Histogram Tool).
4. **Output:** You can choose to save the output of the fit as a file to the specified directory and to subtract the fit from the image and to subtract the fit from the original, creating a **Residual Image** that gets stored as a CASA image and automatically loaded into the viewer. This gives a way to tell how well your fit describes the total emission.
5. **Visualization:** You can toggle whether the fit is displayed on the viewer or not and change the color of the marker.

Click **Fit** to start the fit. If the fit does not converge, try improving your initial estimates and fitting again.

7.4.9 Interactive Position-Velocity Diagram Creation

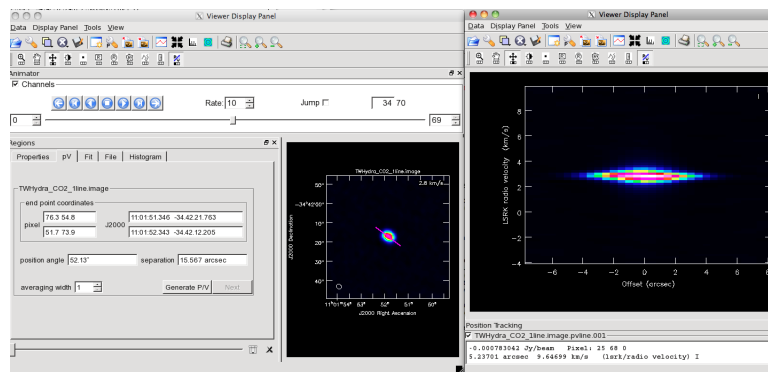


Figure 7.33: Interactive creation of position-velocity cuts in the viewer. Use the P/V tool from the Mouse Toolbar to define a cut, then use the pV tool from the Region Manager Panel to adjust the cut (including the width). Click Generate P/V to build the position velocity cut and open it in a new Viewer Display Panel (from which it can be saved to disk).

The route to create position-velocity cuts in the viewer is illustrated in Figure 7.33:

1. Select the P/V cut tool from the Mouse Toolbar and use it to draw a line across a data cube along the axis you want to visualize.
2. Open the Region Manager Panel and go to the **pV** tab. Highlight the cut you just drew. You should see the end point coordinates listed, along with information on the length and position angle of the cut. You can set the averaging width (in pixels) in a window at the bottom of the tab.
3. When you are satisfied, hit **Generate P/V**. This will create a new Main Viewer Display Panel showing the position velocity cut. The axes should be Offset and velocity.

The new image can be saved to disk with the **Data:Save as...** option.

7.5 Viewing Measurement Sets

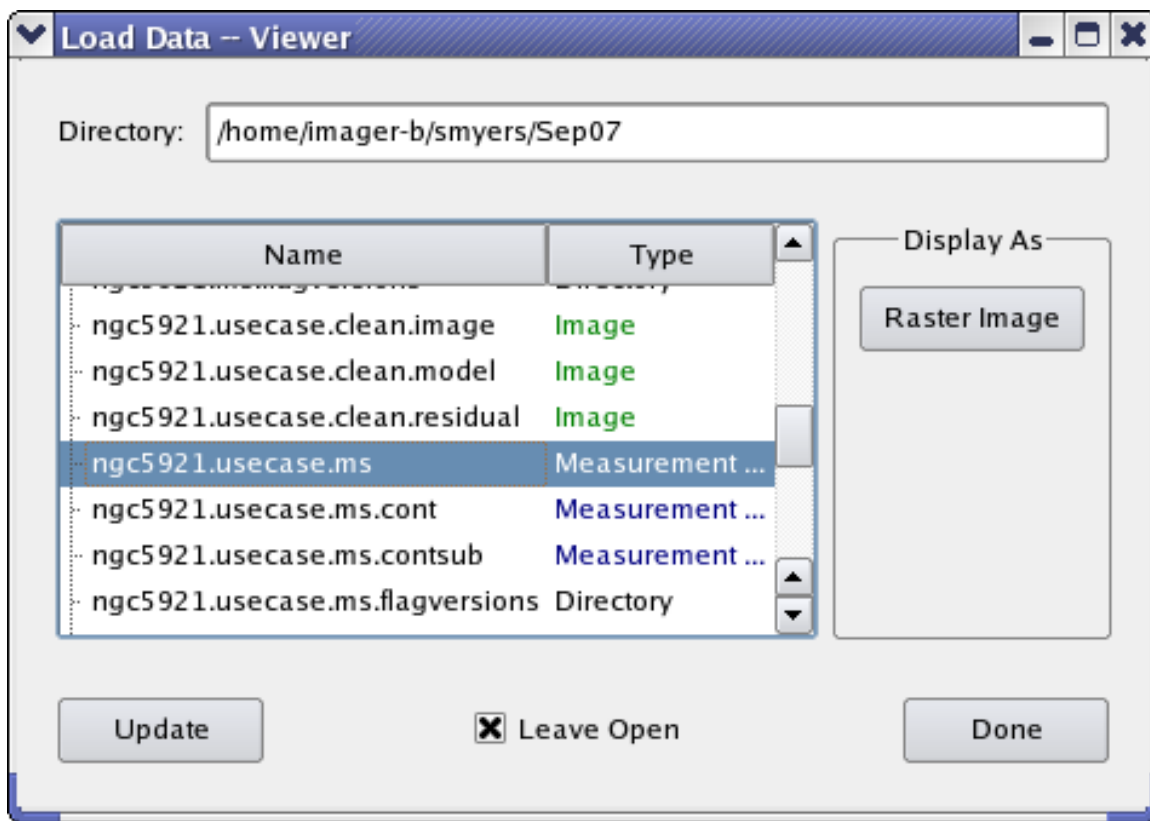


Figure 7.34: The **Load Data - Viewer** panel as it appears if you select an MS. The only option available is to load this as a **Raster Image**. In this example, clicking on the **Raster Image** button would bring up the displays shown in Figure 7.2.

Visibility data can also be displayed and flagged directly from the viewer. For Measurement Set files the only option for display is 'Raster' (similar to AIPS task TVFLG). An example of MS display is shown in Figure 7.2; loading of an MS is shown in Figure 7.34.

Warning: *Only one MS should be registered at a time on a Display Panel.* Only one MS can be shown in any case. You do not have to close other images/MSs, but you should at least 'unregister' them from the Display Panel used for viewing the MS. If you wish to see other images or MSs at the same time, create multiple Display Panel windows.

7.5.1 Data Display Options Panel for Measurement Sets

The **Data Display Options** panel provides adjustments for MSs similar to those for images, and also includes flagging options. As with images, this window appears when you choose the **Data:Adjust** menu or use the wrench icon from the **Main Toolbar**. It is also shown by default when an MS is loaded. The right panel of Figure 7.2 shows a **Data Options** window. It has a tab for each open MS, containing a set of categories. The options within each category can be either 'rolled up' or expanded by clicking the category label.

For a Measurement Set, the categories are:

- **Advanced**
- **MS and Visibility Selection**
- **Display Axes**
- **Flagging Options**
- **Basic Settings**
- **Axis Drawing and Labels**
- **Color Wedge**

7.5.1.1 MS Options — Basic Settings

The **Basic Settings** roll-up is expanded by default. It contains entries similar to those for a raster image (§ 7.4.1.2). Together with the brightness/contrast and colormap adjustment icons on the **Mouse Toolbar** of the Display Panel, they are especially important for adjusting the color display of your MS.

The available Basic options are:

- **Data minimum/maximum**

This has the same usage as for raster images. Lowering the data maximum will help brighten weaker data values.

- **Scaling power cycles**

This has exactly the same usage as for raster images (see § 7.4.1.2). Again, lowering this value often helps make weaker data visible. If you want to view several fields with very different amplitudes simultaneously, this is typically one of the best adjustments to make early, together with the `Colormap` fiddling mouse tool, which is on the middle mouse button by default.

- **Colormap**

Greyscale or Hot Metal colormaps are generally good choices for MS data.

7.5.1.2 MS Options— MS and Visibility Selections

- **Visibility Type**

- **Visibility Component**

- **Moving Average Size**

This roll-up provides choice boxes for Visibility Type (Observed, Corrected, Model, Residual) and Component (Amplitude, Phase, Real, or Imaginary).

Changes to Visibility Type or Component (changing from Phase to Amplitude, for example) require the data to be retrieved again from the disk into memory, which can be a lengthy process. When a large MS is first selected for viewing, the user must trigger this retrieval manually by pressing the **Apply** button (located below all the options), after selecting the data to be viewed (see `Field IDs` and `Spectral Windows`, below).

Tip: Changing visibility type between 'Observed' and 'Corrected' can also be used to assure that data and flags are reloaded from disk. You should do this if you're using another flagging tool such as `autoflag` simultaneously, so that the viewer sees the other tool's new edits and doesn't overwrite them with obsolete flags. The **Apply** button alone won't reload unless something within the viewer itself requires it; in the future, a button will be provided to reload flags from the disk unconditionally.

You can also choose to view the difference from a running mean or the local RMS deviation of either Phase or Amplitude. There is a slider for choosing the nominal number of time slots in the 'local neighborhood' for these displays.

(Note: **Insufficient Data** is shown in the tracking area during these displays when there is no other unflagged data in the local neighborhood to compare to the point in question. The moving time windows will not extend across changes in either field ID or scan number boundaries, so you may see this message if your scan numbers change with every time stamp. An option will be added later to ignore scan boundaries).

- **Field IDs**

- **Spectral Windows**

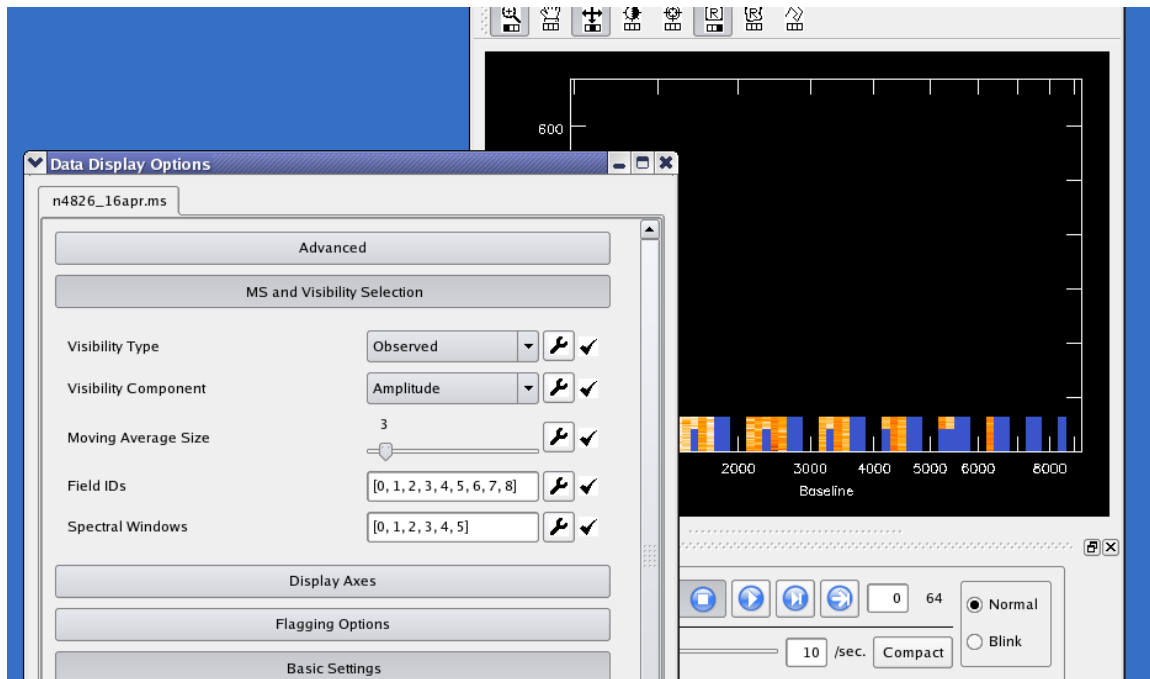


Figure 7.35: The MS for NGC4826 BIMA observations has been loaded into the viewer. We see the first of the `spw` in the Display Panel, and have opened up **MS and Visibility Selections** in the **Data Display Options** panel. The display panel raster is not full of visibilities because `spw 0` is continuum and was only observed for the first few scans. This is a case where the different spectral windows have different numbers of channels also.

You can retrieve and edit a selected portion of the MS data by entering the desired Spectral Window and Field ID numbers into these boxes. **Important:** Especially with large MSs, often the first thing you'll want to do is to select **spectral windows** which all have the **same number of channels** and the **same polarization setup**. It also makes sense to edit only a few fields at a time. Doing this will also greatly reduce data retrieval times and memory requirements.

You can separate the ID numbers with spaces or commas; you do not need to enter enclosing brackets. Changes to either entry box will cause the selected MS data to be reloaded from disk.

If you select, say, spectral windows 7, 8, 23, and 24, the animator, slice position sliders, and axis labeling will show these as 0, 1, 2, and 3 (the 'slice positions' or 'pixel coordinates' of the chosen spectral windows). Looking at the position tracking display is the best way to avoid confusion in such cases. It will show something like: `Sp Win 23 (s 2)` when you are viewing spectral window 23 (plane 2 of the selected spectral windows).

Changes to MS selections will not be allowed until you have saved (or discarded) any previous edits you have made (see **Flagging Options -- Save Edits**, below). A warning is printed on the console (not the logger).

Initially, all fields and spectral windows are selected. To revert to this 'unselected' state, choose 'Original' under the wrench icons next to the entry boxes.

See Figure 7.35 for an example showing the use of the **MS and Visibility Selection** controls when viewing an MS.

7.5.1.3 MS Options — Display Axes

This roll-up is very similar to that for images: it allows the user to choose which axes (from Time, Baseline, Polarization, Channel, and Spectral Window) are on the display and the animator. There are also sliders here for choosing positions on the remaining axes. (It's useful to note that the data *is* actually stored internally in memory as an array with these five axes).

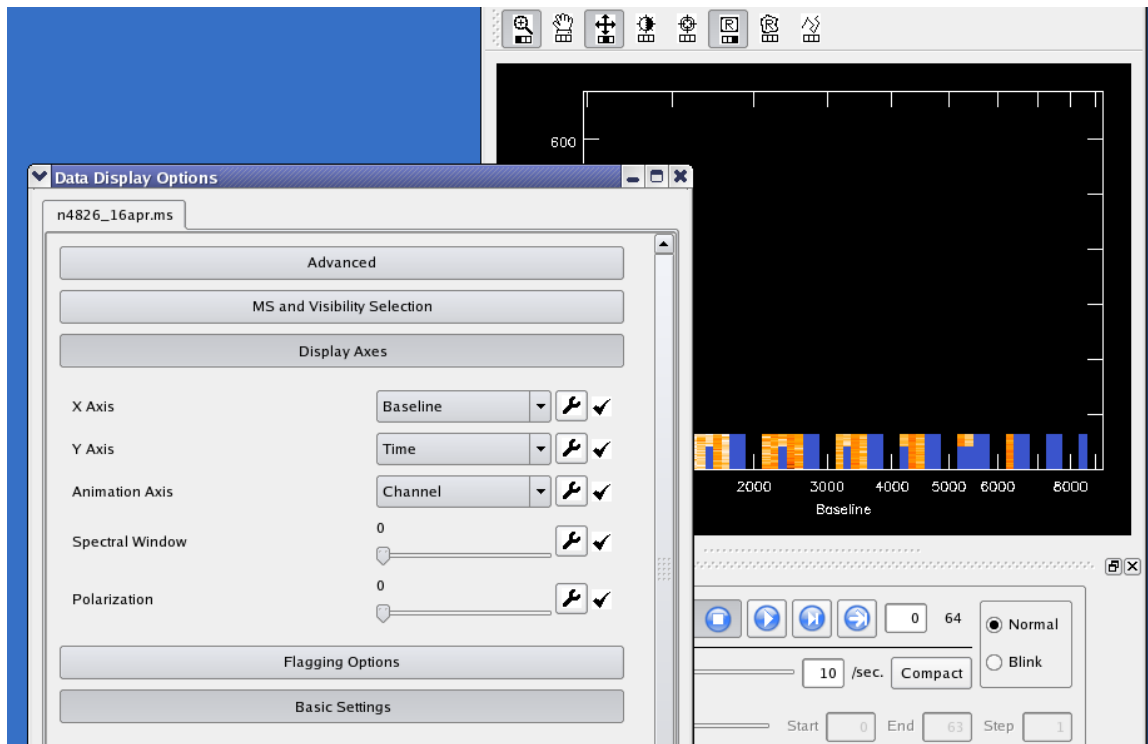


Figure 7.36: The MS for NGC4826 from Figure 7.35, now with the **Display Axes** open in the **Data Display Options** panel. By default, channels are on the **Animation Axis** and thus in the tapedeck, while spectral window and polarization are on the **Display Axes** sliders.

For MSs, changing the choice of axis on one control will automatically swap axes, maintaining different axes on each control. Changing axes or slider/animator positions does not normally require pressing **Apply** — the new slice is shown immediately. However, the display may be partially or completely grey in areas if the required data is not currently in memory, either because no data has been loaded yet, or because not all the selected data will fit into the allowed memory.

Press the **Apply** button in this case to load the data (see § 7.5.1.6 and `Max. Visibility Memory` at the end of § 7.5.1.5).

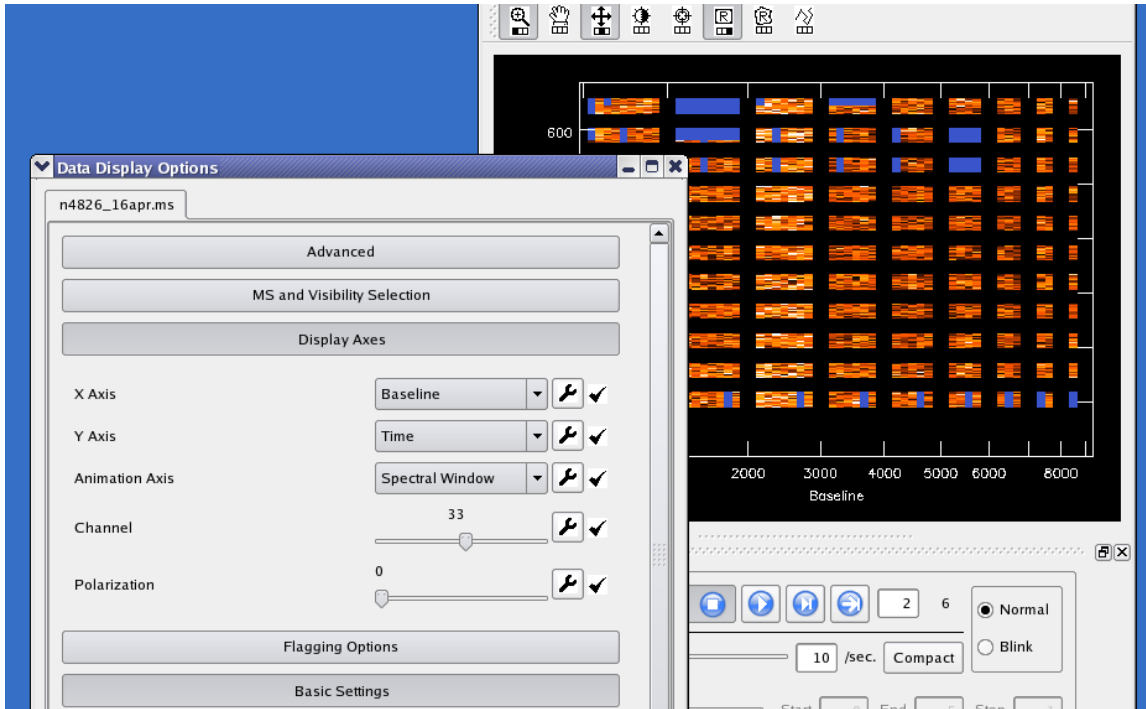


Figure 7.37: The MS for NGC4826, continuing from Figure 7.36. We have now put `spectral window` on the **Animation Axis** and used the tape deck to step to `spw 2`, where we see the data from the rest of the scans. Now `channels` is on a `Display Axes` slider, which has been dragged to show `Channel 33`.

Within the `Display Axes` rollup you may also select whether to order the baseline axis by `antenna1-antenna2` (the default) or by (unprojected) baseline length.

See Figures 7.36–7.37 showing the use of the `Display Axes` controls to change the axes on the animation and sliders.

7.5.1.4 MS Options — Flagging Options

These options allow you to edit (flag or unflag) MS data. The `Point Tool` and `Rectangle Region Mouse Tools` (§ 7.4.3.2) are used on the display to select the area to edit. When using the `Rectangle Region` tool, double-click inside the selected rectangle to confirm the edit.

The options below determine how edits will be applied.

- Show Flagged Regions...

You have the option to display flagged regions in the background color (as in TVFLG) or to highlight them with color. In the former case, flagged regions look just like regions of no data. With the (default) color option, flags are shown in shades of blue: darker blue for flags already saved to disk, lighter blue for new flags not yet saved; regions with no data will be shown in black.

- **Flag or Unflag**

This setting determines whether selected regions will be flagged or unflagged. This does *not* affect previous edits; it only determines the effect which later edits will have. Both flagging and unflagging edits can be accumulated and then saved in one pass through the MS.

- **Flag/Unflag All...**

These flagging extent checkboxes allow you to extend your edit over any of the five data axes. For example, to flag *all* the data in a given time range, you would check all the axes *except* Time, and then select the desired time range with the **Rectangle Region** mouse tool. Such edits will extend along the corresponding axes over the entire selected MS (whether loaded into memory or not) and optionally over unselected portions of the MS as well (**Use Entire MS**, below). Use care in selecting edit extents to assure that you're editing all the data you wish to edit.

- **Flag/Unflag Entire Antenna?**

This control can be used to extend subsequent edits to all baselines which include the desired antenna[s]. For example, if you set this item to 'Yes' and then click the point tool on a visibility position with baseline 3-19, the edit would extend over baselines 0-3, 1-3, 2-3, 3-3, 3-4, ... 3-*n*Antennas-1. Note that the second antenna of the selection (19) is irrelevant here – you can click anywhere within the 'Antenna 3 block', i.e., where the *first* antenna number is 3, to select all baselines which include antenna 3.

This item controls the edit extent only along the baseline axis. If you wish to flag *all* the data for a given antenna, you must still check the boxes to flag all Times, Channels, Polarizations and Spectral Windows. There would be no point, however, in activating *both* this item and the 'Flag All Baselines' checkbox. You can flag an antenna in a limited range of times, etc., by using the appropriate checkboxes and selecting a rectangular region of visibilities with the mouse.

Note: You do not need to include the entire 'antenna block' in your rectangle (and you may stray into the next antenna if you try). Anywhere within the block will work. To flag higher-numbered antennas, it often helps to zoom in.

- **Undo Last Edit**

- **Undo All Edits**

The 'Undo' buttons do the expected thing: completely undo the effect of the last edit (or all unsaved edits). Please note, however, that only unsaved edits can be undone here; there is no ability to revert to the flagging state at the start of the session once flags have been saved to disk (unless you have previously saved a 'flag version'. The flag version tool is not available through the viewer directly).

- **Use Entire MS When Saving Edits?**

”Yes” means that saving the edits will flag/unflag over the entire MS, *including* fields (and possibly spectral windows) which are not currently selected for viewing. Specifically, data within time range(s) you swept out with the mouse (even for unselected fields) will be edited.

In addition, if ”Flag/Unflag All...” boxes were checked, such edits will extend throughout the MS. Note that only unselected *times* (fields) can be edited *without* checking extent boxes for the edits as well. Unselected spectral windows, e.g., will *not* be edited unless the edit also has ”Flag/Unflag All Spectral Windows” checked.

Warning: Beware of checking “All Spectral Windows” unless you have also checked ”All Channels” or turned “Entire MS” off; channel edits appropriate to the selected spectral windows may not be appropriate to unselected ones. Set ”Use Entire MS” to “No” if your edits need to apply only to the portion of the MS you have selected for viewing. *Edits can often be saved significantly faster this way as well.*

Also note that checkboxes apply to individual edits, and must be checked before making the edit with the mouse. “Use Entire MS”, on the other hand, applies to all the edits saved at one time, and must be set as desired before pressing ”Save Edits”.

- **Save Edits**

MS editing works like a text editor in that you see all of your edits immediately, but nothing is committed to disk until you press “Save Edits”. Feel free to experiment with all the other controls; nothing but ’Save Edits’ will alter your MS on disk. As mentioned previously, however, there is no way to undo your edits once they are saved, except by manually entering the reverse edits (or restoring a previously-saved ’flag version’).

Also, *you must save* (or discard) *your edits before changing the MS selections*. If edits are pending, the selection change will not be allowed, and a warning will appear on the console.

If you close the MS in the viewer, *unsaved edits are simply discarded*, without prior warning. It’s important, therefore, to remember to save them yourself. You can distinguish unsaved flags (when using the ’Flags In Color’ option), because they are in a lighter shade of blue.

The program must make a pass through the MS on disk to save the edits. This can take a little time; progress is shown in the console window.

7.5.1.5 MS Options— Advanced

These settings can help optimize your memory usage, especially for large MSs. A rule of thumb is that they can be increased until response becomes sluggish, when they should be backed down again.

You can run the unix ’top’ program and hit ’M’ in it (to sort by memory usage) in order to examine the effects of these settings. Look at the amount of RSS (main memory) and SWAP used by the X server and ’casaviewer’ processes. If that sounds familiar and easy, then fiddling with these settings is for you. Otherwise, the default settings should provide reasonable performance in most cases.

- **Cache size**

The value of this option specifies the maximum number of different views of the data to save so that they can be redrawn quickly. If you run an animation or scroll around zoomed data, you will notice that the data displays noticeably faster the second time through because of this feature. Often, setting this value to the number of animation frames is ideal. Note, however, that on multi-panel displays, each panel counts as one cached image.

Large images naturally take more room than small ones. The memory used for these images will show up in the X server process. If you need more Visibility Memory (below) for a really large ms, it is usually better to forgo caching a large number of views.

- **Max. Visibility Memory**

This option specifies how many megabytes of memory may be used to store visibility data from the measurement set internally. *Even if you do not adjust this entry, it is useful to look at it to see how many megabytes are required to store your entire (selected) MS in memory.* If the slider setting is above this, the whole selected MS will fit into the memory buffer. Otherwise, some data planes will be 'grayed out' (see **Apply Button**, § 7.5.1.6 below), and the selected data will have to be viewed one buffer at a time, which is somewhat less convenient. In most cases, this means you should **select fewer fields or spectral windows** – see § 7.5.1.2. The 'casaviewer' process contains this buffer memory (it contains the entire viewer, but the memory buffer can take most of the space).

7.5.1.6 MS Options — Apply Button

When viewing large MSs the display may be partially or completely grey in areas where the required data is not currently in memory, either because no data has been loaded yet, or because not all the selected data will fit into the allowed memory (see **Max. Visibility Memory** above). When the cursor is over such an area, the following message shows in the position tracking area:

```
press 'Apply' on Adjust panel to load data
```

Pressing the **Apply** button (which lies below all the options) will reload the memory buffer so that it includes the slice you are trying to view.

The message **No Data** has a different meaning; in that case, there simply *is* no data in the selected MS at the indicated position.

For large measurement sets, loading visibility data into memory is the most time-consuming step. Progress feedback is provided in the console window. Again, careful selection of the data to be viewed can greatly speed up retrieval.

7.6 Printing from the Viewer

You can select **Data:Print** from the drop down menu or click the **Print** icon to bring up the **Viewer Print Manager**. From this panel, you can **Print** the contents of Display Panel to a

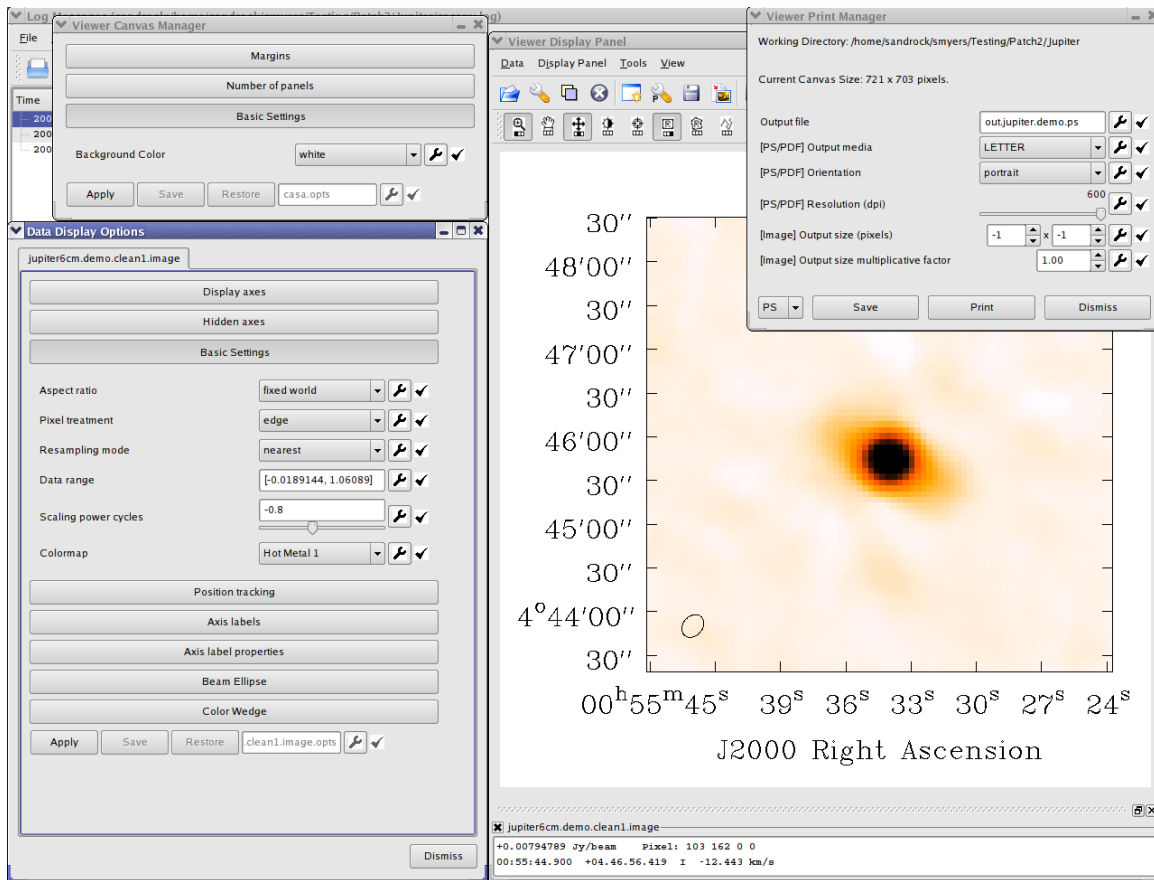


Figure 7.38: Printing the display to a hardcopy of a file. From the **Viewer Print Manager**, located in top right here and accessed by the print icon or from the Data drop down menu, you can use the **Save** button to save an image or **Print** directly to a printer. To achieve the best results, it is often helpful to adjust the settings in the **Data Display Options** and **Viewer Canvas Manager**, shown at right.

hardcopy or **Save** them as an image in a format selected from the drop-down menu at the bottom left of the window. Note that the save feature will overwrite the file in question without prompting.

The Viewer Print Manager allows you to adjust the DPI, orientation, and page format (Output Media) for Postscript or PDF files and to scale the image to a desired pixel size for other images.

To achieve the best output it is usually advisable to adjust the settings in the **Viewer Print Manager**, **Data Display Options**, and **Viewer Canvas Manager**. For PDF and Postscript output, turning the DPI up all the way yields the best-looking results. For other images, a white background often makes for better looking images than the default black. It is often necessary to increase the **Line Width** in the **Axis Label Properties** (in the **Data Display Options** panel) to ensure that the labels will be visible when printed. Increasing from the default of 1.4 to a value around 2 often works well.


```

out          =          ''          # (Optional) Output filename or
                                     # complete output config dictionary.
                                     # If a string is passed, the file
                                     # extension is used to determine the
                                     # output type (jpg, pdf, eps, ps, png,
                                     # xbm, xpm, or ppm). If a dictionary
                                     # is passed, it can contain the
                                     # fields, file (string), scale
                                     # (float), dpi (int), or orient
                                     # (landscape or portrait). The scale
                                     # field is used for the bitmap formats
                                     # (i.e. not ps or pdf) and the dpi
                                     # parameter is used for scalable
                                     # formats (pdf or ps).

```

The `raster` and `contour` parameters specify which images to load and how these images should be displayed. These parameters take python dictionaries as inputs. The fields in these dictionaries specify how the image will be displayed.

An example call to `imview` looks like this:

```

imview(raster={'file': 'ngc5921.clean.image',
              'range': [-0.01,0.03],
              'colormap': 'Hot Metal 2',
              'scaling': -1},
       contour={'file': 'ngc5921.clean.image'},
       axes={'x':'Declination'} ,
       zoom={'channel': 7, 'blc': [75,75], 'trc': [175,175],
            'coord': 'pixel'},
       out='myout.png')

```

The argument to `raster` is enclosed in the curly braces `{ }`. Within these braces are a number of "key": "value" pairs. Each sets an option in the viewer, with the GUI parameter to set defined by the "key" and the value to set it to defined by "value." In the example above, `'file':ngc5921.clean.image'` sets the file name of the raster image, `'range': [-0.01,0.03]` sets the range of pixel values used for the scaling.

`contour` works similar to `'raster'` but can accept multiple dictionaries in order to produce multiple contour overlays on a single image. To specify multiple contour overlays, simply pass multiple dictionaries (comma delimited) in to the `contour` argument:

```

contour={'file': 'file1.image', 'levels': [1,2,3] },
        {'file': 'file2.image', 'levels': [0.006, 0.008, 0.010] }

```

`zoom` specifies the part of the image to be shown.

`axes` defines what axes are shown. By default, the viewer will show `'x': 'Right Ascension', 'y': 'Declination'` but one may also view position-frequency images.

out defines the filename of the output, with the extension setting the file type.

Currently, the following parameters are supported with additional functionality planned for future releases:

```
raster -- (string) image file to open
        (dict)  file (string)    => image file to open
                scaling (float)  => scaling power cycles
                range (float*2)  => data range
                colormap (string) => name of colormap
                colorwedge (bool) => show color wedge?
contour -- (string) file to load as a contour
        (dict)  file (string)    => file to load
                levels (float*N) => relative levels
                base (numeric)   => zero in relative levels
                unit (numeric)   => one in the relative levels
zoom    -- (int)    integral zoom level
        (string)  region file to load as the zoom region
        (dict)    blc (numeric*2) => bottom left corner
                 trc (numeric*2) => top right corner
                 coord (string)  => pixel or world
                 channel (int)    => channel to display
        (dict)    <region record> => record loaded
                 e.g. rg.fromfileto record( )
axes    -- (string*3) dimension to display on the x, y, and z axes
        (dict)    x              => dimension for x-axes
                 y              => dimension for y-axes
                 z              => dimension for z-axes
out     -- (string) file with a supported extension
        [jpg, pdf, eps, ps, png, xbm, xpm, ppm]
        (dict)    file (string)  => filename
                 format (string) => valid ext (filename ext overrides)
                 scale (numeric) => scale for non-eps, non-ps output
                 dpi (numeric)   => dpi for eps or ps output
                 orient (string) => portrait or landscape
```

Examples are also found in `help imview`.

7.8 Measurement Viewer (msview)

The Measurement Viewer `msview` is mostly a clone of the `viewer` at this stage. A difference is that `msview` allows the user to select data before it is loaded into the GUI and displayed. A screenshot is shown in Fig. 7.39 and selection parameters are `field`, `spectral window`, `time range`, `uv range`, `antenna`, `corr`, `scan`, `array`, `ms selection expression` in the usual CASA selection syntax (see Sect. 2.3).

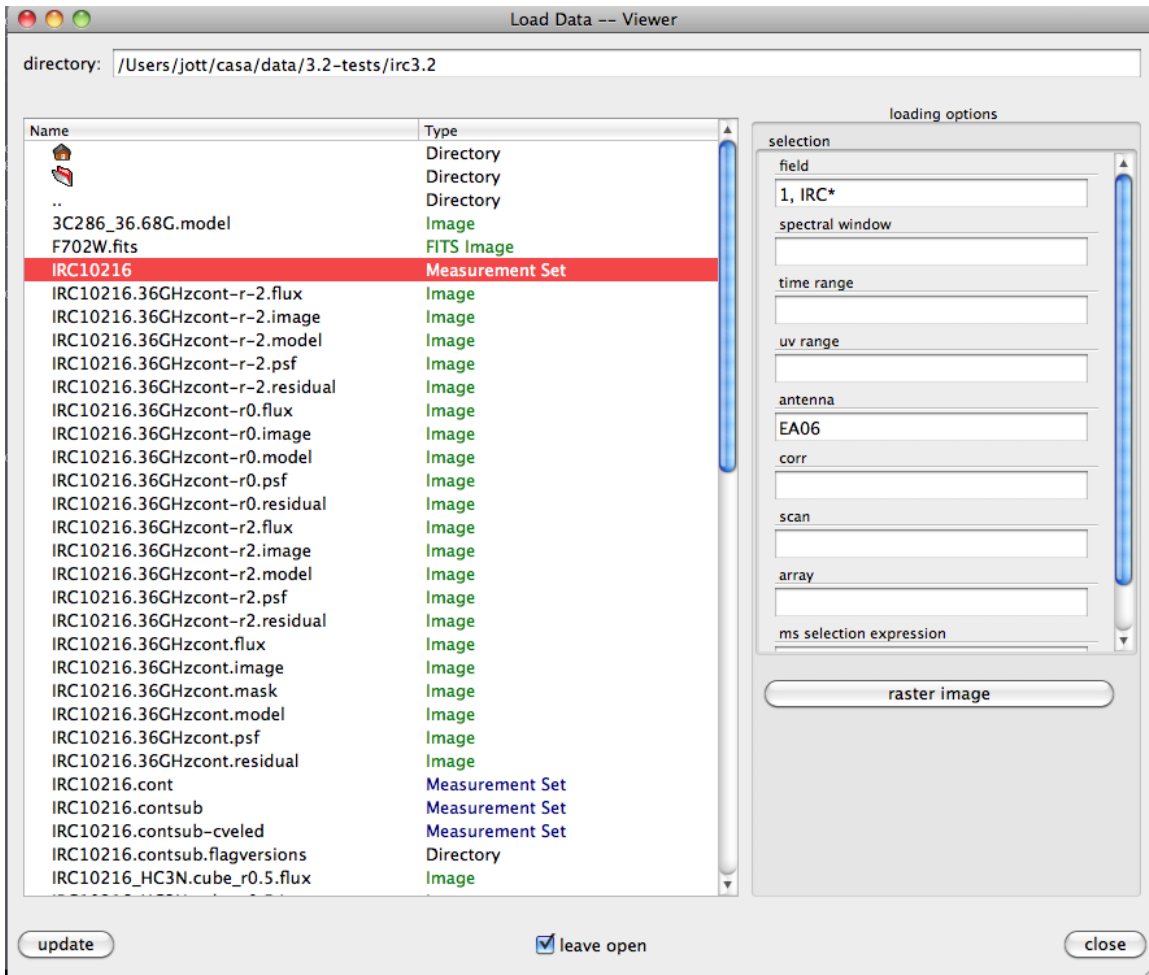


Figure 7.39: Data selection in msview.

Chapter 8

Single-dish Data Processing

This chapter deals with a description of the single-dish toolset of CASA. The target of CASA single-dish development is to calibrate, reduce and image data taken with both the ALMA single-dish telescopes and the Nobeyama single-dish telescope. The CASA single-dish tool suite is built separately to the interferometer suite, although with a growing overlap of base code and toolsets. Primarily this is because the single-dish part of CASA has historically made use the ATNF Spectral Analysis Package (ASAP) to expedite development, although CASA SD tools are now moving towards a structure more closely associated with the interferometer toolset, which use "Measurement Set" format. In CASA 4.5, the processing of ALMA single-dish data can be made completely in Measurement Set format. The migration to the remaining (non-core) CASA SD toolset, including analysis tools more "advanced" reduction tools, will continue through the next CASA versions. Importantly, ASAP is still supported in CASA 4.5 although it will be phased out in coming versions. In the cases ASAP format is used in any tools, the format conversion is intrinsically imported within the tool, and any calls are invoked transparently to the operator.

For details on ASAP – including the User Guide, Reference Manual, and tutorial – see the ASAP home page at ATNF:

<http://svn.atnf.csiro.au/trac/asap/>.

In this chapter, we detail the current status of CASA SD implementation and functionality. Covering CASA SD-specific environments, recent progress in design development.

For trouble-shooting, please be sure to check the list of known issues and features of ASAP and the SDtasks presented in Sect. 1.2.1 and http://casa.nrao.edu/release_ki.shtml first.

8.1 Current status, new tasks/operations and CASA SD setup

8.1.1 Transition from ASAP *Scantable* format, and using *Scantable* format

All ALMA raw data, including data from the Total Power (TP) array of ALMA are in the *ALMA Science Data Model* (ASDM) format. Prior to CASA 4.4, Total Power ASDMs were processed using

ASAP toolkit, which required data conversion to *Scantable* format. As processing of interferometric data is done in *Measurement Set* it is more practical from a software development and maintenance point of view for the Total-power aspect of CASA to also employ *Measurement Set* format. Even so, for CASA 4.5, we continue to support processing in *Scantable* format - however *support for Scantable format will be phased out over the following few CASA versions.*

A number of environment variables used by CASA SD tools (see `.asaprc`). Within CASA, they are accessible through keys and values of the Python dictionary `sd.rcParams`. e.g:
Set verbose mode (producing verbose feedback form CASA SD).

```
sd.rcParams['verbose'] = True
```

Set so *Scantable* operations are done in memory or on disk (use 'disk' when the dataset is large.)

```
sd.rcParams['scantable.storage'] = 'memory' (default)
```

(Note: setting `sd.rcParams['scantable.storage'] = 'disk'` may cause data to be overwritten when using some tool and functions, even if `sd.rcParams['insitue'] = False`). See § 8.7 for more details on ASAP-based CASA tasks which affect on-disk data. See § 8.1.1 for more details on the ASAP environment variables.

8.1.2 Dictionaries

Many SDtasks return Python dictionaries. E.g. The results of line fitting (in `sdfit`) and region statistics (in `sdstat`). The relevant parameters returned by the tasks can be extracted using the appropriate keys:e.g.,

```
CASA <10>: line_stat=sdstat()
Current fluxunit = K
No need to convert fluxunits
Using current frequency frame
Using current doppler convention
```

```
CASA <11>: line_stat
Out[11]:
{'eqw': 70.861755476162784,
 'max': 1.2750182151794434,
 'mean': 0.35996028780937195,
 'median': 0.23074722290039062,
 'min': -0.20840644836425781,
 'rms': 0.53090775012969971,
 'stddev': 0.39102539420127869,
 'sum': 90.350028991699219}
```

One can then use these values in scripts by accessing this dictionary, e.g.,

```
CASA <12>: print "Line max = %5.3f K" % (line_stat['max'])
Line max = 1.275 K
```


8.1.3 Line Formatting

CASA tasks operating on Measurement Set format are able to trap leading and trailing whitespace on string parameters (such as `infile`) but ASAP does not. Note also; ASAP is case-sensitive, with most parameters being upper-case (e.g. 'ASAP' used in `sd.scantable.save`)

8.1.4 New Tasks, Features and functionality in CASA 4.5

As for CASA 4.4, development towards CASA 4.5 release has been dominated by the porting to CASA Measurement Set format, without prioritizing development of new tasks. In this way the user experience for processing single-dish data is more consistent with that for processing interferometer data. A complete list of Measurement Set -capable tasks is at the end of this chapter.

The issue of non-linearity in the ALMA correlator is now solved at the point of data-taking. Single-dish data taken before cycle 3 will still suffer from that problem, and will need to be properly treated in processing (see section 8.2.1).

As the effort is largely diverted towards completing the transition to use Measurement Set format, there are smaller improvements to existing functionality - such as the selection by observation intent in the imaging task `sdimaging` (i.e. select using the 'purpose' for a particular observation: target, or reference, etc.), streamlining the feedback for the baselining task `tsdbaseline`.

To a very great extent, CASA SD shares even more tasks with CASA interferometer. These include displaying data contents (`listobs`), computing and applying calibration tables (`gencal`, `applycal`), data flagging and subselection (`flagdata`, `mstransform`) and moment image generation (`imcontsub`, `immoment`; see Section 8.5). The drive towards merging the 'single-dish side' and 'interferometer side' of CASA is ongoing.

8.2 SD data-taking in brief

As for interferometer data (see Chapter 2), single-dish data from ALMA is initially in a format called the ALMA-Science Data Model (ASDM). These ASDMs must be converted at least to Measurement Set for many single-dish tasks (Some CASA single-dish tasks also requires conversion to *Scantable* format, but this happens transparently to the operator).

Single-dish observations taken with ALMA include the science target (obviously) along with observations of a quasar (for band 3) or Uranus. These quasar/Uranus observations are observed differently, and processed separately to the science data, and are used to provide an empirical calibration to scale from Kelvin to Janskys, as shown in Figures 8.1.

It is an ALMA policy that the measurements of the Jy/K calibrator be taken within a few days of the science observations. Both the calibration and science observations also have their own pointing, focus, T_{sys} and reference (short-term calibration) observations in the following order - with the noted exception:

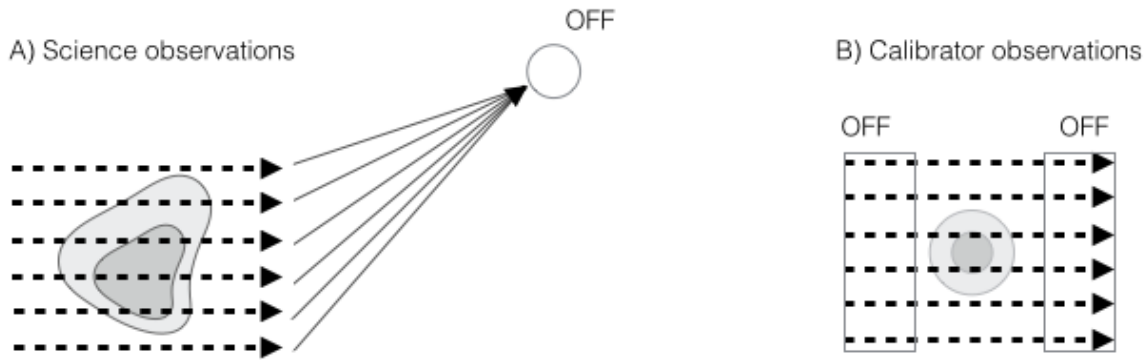


Figure 8.1: Observations modes for science observations *A) Left* where the OFF position is distant from the target field, and calibrator observations *B), Right* where the OFF positions are extracted from the edges of the scanned area.

wh

1. Focus optimisation (data discarded from data reduction)
2. Pointing optimisation (data discarded from data reduction)
3. T_{sys} calibration
4. Pointing Measurement of Reference (OFF) **science observations only, not calibrator**
5. T_{sys} calibration
6. OTF Measurement of source field.

The OFF position is not observed during the observations of the calibrator (quasar/Uranus), because those observations are fast enough relative to the atmospheric stability, and those data are calibrated relative to "edge of map" data, rather than explicit OFF data.

8.2.1 Correlator non-linearity

In 2014 it was determined ASDM products of the ALMA correlator must be scaled to compensate for nonlinear terms, and that the adjustments can be made at the point of data-taking - i.e. during the observations. From cycle 3 and onwards, the scaling is done at the point of data-taking (i.e. on-the-fly calibration), however data taken before Cycle 3 should be checked to see if it is necessary to scale the data in CASA. The adjustment to compensate for the non-linearity is effected simply by multiplying the calibrated data by a factor of 1.25. This can be achieved in CASA 4.5 using the `gencal` and `applycal` tasks.

8.3 Overview of SD tools and tasks

In the following, we describe the *core* single-dish tasks, used most frequently in the contexts of reduction tools and analysis tools. Non-core tasks (i.e. tasks non-essential for basic data reduction, image generation and analysis) as well as tasks that are expected to be depreciated in the near future are briefly described in Section 8.5, and the remainder of the entire single-dish ASAP-based SD tool set are detailed in Section 8.7.

Core CASA reduction task descriptions

- **tsdcal** — generate sky and T_{sys} caltables for SD data, and apply them.
- **sdimaging** — create an image from the total power or spectral data.
- **tsdbaseline** — compute fit for per-spectrum baseline and optionally removes it.

Core CASA SD analysis task descriptions

- **sdaverage** — average, smooth and regrid SD spectra.
- **tsdfit** — line fitting to SD spectra.
- **sdmath** — do simple arithmetic for SD datasets.
- **sdscale** — scale SD data with a given factor.
- **sdstat** — compute statistics of regions of SD spectra.
- **tsdsmooth** — gaussian smoothing along spectral axis.

Other CASA tasks relevant for SD data reduction and analysis

- **importasdm** — Convert an ALMA Science Data Model (ASDM) into a visibility file (MS).
- **listobs** — List the summary of a data set in the logger or in a file.
- **flagdata** — All-purpose flagging task based on data-selections and flagging modes.
- **split** — Create a visibility subset from an existing visibility set.
- **plotms** — A plotter/interactive flagger for visibility data.
- **genical** — Specify Calibration Values of Various Types.
- **applycal** — Apply calibrations solutions(s) to data.
- **imhead** — List, get and put image header parameters.
- **imcontsub** — Estimates and subtracts continuum emission from an image cube.
- **immoments** — Compute moments from an image.

8.4 Overview for Reducing ALMA Single-dish Data

As described in Section 8.2, Single-dish science data taken in Cycle 3 are calibrated with separate observations of a quasar (band 3 observations), or Uranus (all other bands). Both the science and calibrator datasets are separately processed (i.e. they have their own measurements of T_{sys} separately applied). The Calibrator observations are compared with ALMA databases to determine a factor to convert the native Kelvin units of the single-dish observations, into the Janskys units of the measurements held in the ALMA database. This conversion step is necessary, to properly scale for combination of the single-dish and interferometer datasets.

Therefore, the Calibrator observations should be conducted first, although they differ only in the application of a "baseline removal" step and the parameters of the image product (i.e. the calibrator observations yield a 2-D continuum image, whereas the science observations yield a 3-D position-position-velocity cube).

The workflow is shown in Figure 8.2, and each step is detailed below. The tasks applied for the reduction are described in following sections which are organised by coarsely grouping the tasks into levels of functionality.

The `Analysis_Uutilities` package will be used only during the imaging steps. The package can be imported, and the `'stuffForScienceDataReduction'` class can be instantiated with:

```
import analysisUtils as aU
es = aU.stuffForScienceDataReduction()
```

Please access the `Analysis_Uutils` package from this link: https://casaguides.nrao.edu/index.php/Analysis_Uutilities

8.4.1 Brief Description of SD reduction

The following briefs each of the steps towards fully calibrated image datasets, as shown in Figure 8.2. Note that this process occurs entirely in Measurement Set format.

1. `importASDM`, `import ASDM flags`

The conversion from ASDM to Measurement Set is done with the task `importasdm` (See section 2.2.1). This task also imports various flags and other metadata unique to single-dish, from the ASDM to the measurement set. This is accomplished with the `bdflags2MS` keyword.

2. Inspection of data and T_{sys} quality

To obtain a kind of overview, and determine spectral window identification numbers, etc. we use the `listobs` task. This task has superseded the sd-based task `sdlist`. To examine the quality of the T_{sys} (used to calibrate the data to the kelvin scale), we use the task `gencal` to extract the T_{sys} data as a calibration table, then use `plotbandpass` to examine the stability.

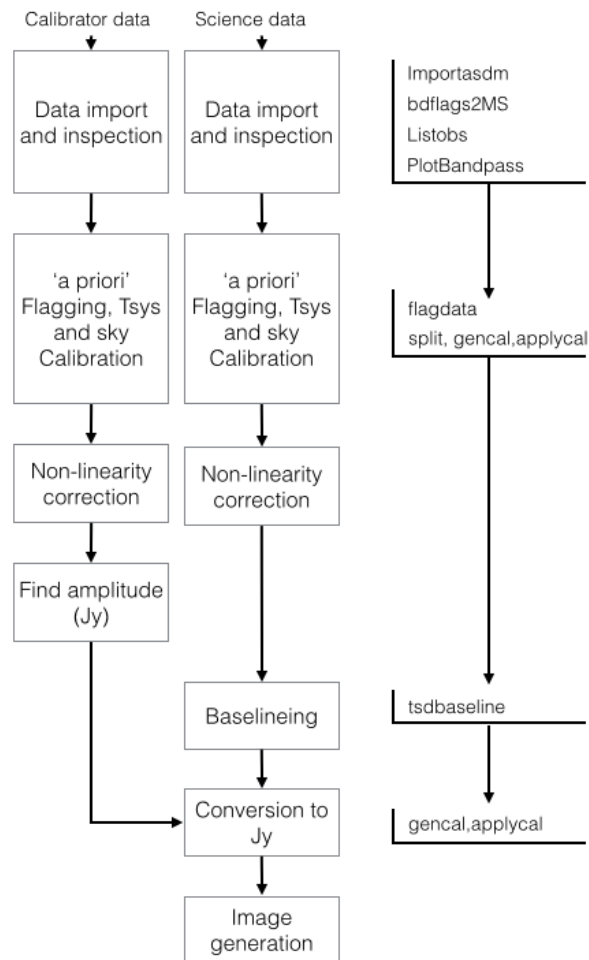


Figure 8.2: General steps for processing both *calibrator* and *science* data, and applying calibrations to science data. The right-most column describes the CASA 4.5 tasks called during each stage.

3. 'a priori' flagging (band edge, any atmospheric lines)

Typically, the edges of the bandpasses dive to zero power. This is a symptom of all tuned radio telescopes and is simply a manifestation of the frequency-sampling function across the baseband. This step is only important if the spectral window also samples the edge of the baseband and in this case we typically flag a few percent of each band - approximately 5-10% of the affected edges of the spectral windows using the `flagdata` task (which has replaced the task `sdflag`). The unflagged data should be examined with `plotms` to ensure the edges of the bands are removed.

The case where a spectral line of scientific value falls close to the edge of the band, will omit this step of course, and will need to adopt a more careful baseline correction effort at later stages.

4. Applying T_{sys} calibration

Applying the T_{sys} calibration is effected using the `tsdcal` task (replacing the `sdcal` and `sdcal2` tasks).

Ostensibly, we calibrate the data into brightness temperature in units of K, using the equation

$$T_{A^*} = T_{\text{sys}} \times (ON - OFF)/OFF \quad (8.1)$$

However note that additional calibrations need to be applied (described below), before the data are correctly represented in T_{A^*} units.

where ON and OFF are the data on-source (i.e., during the raster scanning) and off-source (on a reference position where only background emission exists), respectively.

The most difficult part of this step is matching the "science" SPW's and T_{sys} SPW's. We determine this with the output from `listobs` task, executed above, however we can use the function `Tsysspwmap` to automatically map T_{sys} SPW's to science SPW's, and pass the results of that function into `tsdcal`.

5. Correction for Correlator non-linearity (only data taken during, or before Cycle 3)

This step can be done at any stage after calibration. It's included at this point because it is a step common to both the calibrator and science datasets. The correction is applied by passing the correction value (simply a factor of 1.25), via a `gencal` call, which generates the correction as a calibration table, and a `applycal` call, which applies the calibration. This is simplified using some basic python scripting, but it's worth noting that the `gencal` call has some internal mathematics that actually applies the square of the *inverse* supplied calibration value - therefore the value passed into `gencal` must be the *inverse square root* of the desired value: i.e. $\frac{1}{x^2}$.

6. Baseline subtraction

Significant work has gone into improving the functionality of `tsdbaseline`, used to compute and apply a residual *baseline* feature in the calibrated spectra. Note that '*baseline*' is typical parlance in single-dish processing, referring to what is commonly low-order structure in the frequency domain. The structure occurs primarily from the changing, turbulent and wet atmosphere in the beam of the single-dish, and is therefore different for each single-dish telescope. It is typical to attempt to remove a rather low-order polynomial or spline. Band 3 is almost immune to atmospheric fluctuations and so orders 1 or 2 are sufficient at those bands. Higher orders are likely for higher bands.

After this stage is complete, the data will be calibrated in units of Kelvin, in antenna temperature units: T_{A^*}

7. Imaging of Calibration data - calibration data only

The imaging step is quite involved, and will be described in the last stage below. The difference in imaging the calibrator is that a single continuum image is generated - in contrast to the science data where a spectral *cube* is created. Note that this step is not critical for generating the science spectral data. It is only important if the single-dish data is to be combined with interferometer data, where it must be converted into Janskys, for correct combination.

8. Determination of K to Jy conversion - calibration data only

After imaging the calibration data, the peak brightness in the calibration image (in antenna temperature, in kelvin) is extracted using `imstat`. The process differs a little if the calibrator is a planet or a quasar, in that the planet observations must also be deconvolved (since the planet is a significant size relative to the beam). To determine the *JperK* value, the known flux density (in Jy) is compared with the extracted peak brightness temperature (in K), and a scaling factor determined simply with the ratio `lux (Jy) /Brightness (K)`. This value is used to scale the science data in the next step.

9. Application of K to Jy conversion - science data only

As for the application of the nonlinearity correction, the conversion factor for Kelvin to Jy is simply multiplication by a factor. Again, we pass the correction factor via a `gencal` call and a `applycal` call, to generate and apply the calibration tables, respectively.

10. Imaging of Science data

For this step (and also for the step to image the calibrator, above), the analysis utils package needs to be loaded.

Many of the parameters are not yet automatically determined by default, and so additional tasks must be called (spatial sampling rates are determined with `aU.getTPSampling`, Beam FWHM is determined with `aU.PrimaryBeamArcsec`) to allocate them to variables which will be passed into the main task: `sdimaging`. Note that `sdimaging` can process multiple *calibrated* datasets simultaneously, and there is no need to split the data into respective antennas. That is: multiple *calibrated* datasets, with multiple single-dish antennas can be processed transparently with `sdimaging`.

The metadata associated with the resulting image must also have the correct flux units written in (`imhead`), and if the data are to be used to combine with interferometer data, the beam parameters must also be correctly re-written (`aU.sfBeam`).

Further baselining is possible with `imcontsub`, and moment images can be generated from the resulting calibrated data, with `immoments`,

8.5 Brief Description of functionality for relevant SD tasks

All of the tasks that operate on measurement set format work from a file on disk rather than from a scantable in memory as the ASAP toolkit does (see § 8.7). When operating on *Scantable* format, the tasks transparently invoke a call to `sd.scantable` to read the data. The scantable objects do not persist within CASA after completion of the tasks and are destroyed to free up memory.

Although the Measurement Set can store data from multiple antennas even if it consists of only single-dish spectra (auto-correlation data), the scantable cannot distinguish data from multiple antennas. It causes a problem when the user processes the Measurement Set using tasks that expect ASAP format. Therefore, when using tasks that can operate also on *Scantable* format to process measurement set format, the id or name of the antenna that the user wants to process must be explicitly specified. This can be done by `antenna` parameter. By default (`antenna=0`), data

associate with antenna id 0 is imported. The `antenna` parameter takes no effect for other input data formats.

In the following, we describe tasks that are exclusive to CASA single-dish. The discussion of other tasks that are common to interferometer processing (e.g. `flagmanager`, `importASDM`) will not be repeated here.

8.5.1 Gridding and imaging:

`sdimaging`, `sdtipimaging`, `sdgrid`, `sdimprocess`

Gridding data follows a few steps: firstly a grid with appropriate parameters is characterised after consultation with metadata and user-inputs, and the grid is then iteratively populated with values extracted and weighted from the actual data. The user has significant control over the various gridding functions, and should bear in mind the resulting resolution is necessarily a convolution of a number of sampling functions (the intrinsic sampling rate across the sky, the telescope beam sampling function, the grid's gridding function and the kernel weighting function). A number of kernels are available in `sdimaging` and `sdgrid`: Primary Beam (PB), Spherical (SF), Box-car (Box), Gaussian (Gauss), and Gaussian×Jinc:

$$Jinc(x) = \frac{J_1(\pi \times x/c)}{(\pi \times x/c)} \quad (8.2)$$

with a first order Bessel function: J_1

The different functions are somewhat (but not completely) akin to the `robust` parameter in interferometric processing, where the weights are distributed across the function to optimise sensitivity, or resolution. The actual shape of the gridding kernel is fully constrained. In combination with kernel truncation scales specified with `convsupport` for spherical kernels, `truncate` for Gaussian and GJinc kernels, and in combination with kernel width scales, specified with `gwidth` for Gaussian and GJinc kernels, and with the jinc function parameter specified with `jwidth`.

The user would be advised to start with either a Spherical function (SF) or Primary Beam (PB) function, and determine on their own basis, which functions are most appropriate for their purpose and by comparison of the resulting RMS (obtained with `imstat`) to the theoretical value.

When building the calibrator continuum image, use `sdimaging` with `mode='channel'`, `nchan=1` and `width=nchan`, where `nchan` is equal to the number of channels in the spectral window (which is shown in the `listobs` command).

The `phasecenter` parameter should be set to the centre of the image, this is automatically determined if left unset.

Please refer to the section on 'Dataset mathematics and manipulation' below, for more information about averaging and weighting parameters.

Task `sdtipimaging` invokes a more fundamental application of gridding, and also attempts some baseline subtraction. This task operates on single antennas, and is largely redundant with the development of `sdimaging`, `tsdbaseline` and `imcontsub`.

Task `sdimprocess` is not a gridding task, but is an image–optimisation task that attempts a fourier-based approach to removing scanning noise by either substituting fourier components from data obtained over a variety of scanning directions (Emerson & Grave 1988), or extrapolation in the fourier domain if only one scanning direction is used (Sofue & Reich 1979). It operates on either CASA or FITS images and returns data in Measurement Set format.

8.5.2 Dataset mathematics and manipulation:

`sdcoadd`, `sdaverage`, `tsdsmooth`, `sdmath`, `sdscale`, `sdcoadd`

It is anticipated that many, or all of these tasks will be unsupported in the near future, since the migration to measurement set format will make them redundant, and ongoing development has enabled other alternatives.

Task `sdcoadd` operates on both Measurement set and *scantable* format data, and simply returns a single, format-free dataset from multiple input scantables. Some tolerances have to be applied when attempting to keep the spectral windows in line, in the frequency domain. In the migration to measurement set format, it is redundant with task `imcat`, but there are few occasions, if any, where it need be invoked standard measurement set reduction.

Task `concat` achieves a similar function to `sdcoadd`.

Task `sdaverage` can be used to average and smooth (but not calibrate) data, instead of the averaging functions implicit in

`tt tsdcal`. Note that due to the operation of the underlying ASAP libraries, averaging by polarization (`polaverage = True`), automatically averages in time too. There is no way to avoid this in `sdaverage`, and the workflow will be to leave polarization averaging until the last step, and apply it during gridding. Smoothing kernels include `hanning`, `gaussian`, `boxcar` and `regrid`.

The weighting parameters of averaging (`'tintsys'`, `'tsys'`, `'tint'`, `'var'` or `'median'`) are attempts to normalise the noise level between accumulated spectra.

From the radiometer equation, the variance in a signal is a function of the system temperature, channel width and integration time:

$$\text{Var}(T_{\text{sys}}, d\nu, dT) = \frac{T_{\text{sys}}^2}{(d\nu \times dT)} \quad (8.3)$$

Where T_{sys} is the measured System temperature in K, $d\nu$ here being the channel width in frequency and dT is the integration time in seconds.

When averaging integrations with say, different T_{sys} , $d\nu$ and/or dT , normalising the data correctly is important. For data where the channel widths are identical, but the integration times are different,

the averaged spectra are found with:

$$T_{av}(\nu) = (T_1(\nu) \times w_1 + T_2(\nu) \times w_2) / (w_1 + w_2) \quad (8.4)$$

where the ratio of weighting parameters, $w_1:w_2$ is found with the ratio of integration times, $dT_1:dT_2$.

The other weighting schemes ('tsys', 'tintsys' are analogues of the 'tint' weighting, but instead capture variation of T_{sys} . eg. normalising under T_{sys} weighting uses $1/T_{sys}^2$.

The alternatives CASA SD considers, and some vague use cases are:

1. normalise by integration time and system temperature (tintsys): $\frac{dT}{T_{sys}^2}$, when $d\nu$ is known to be invariant.
2. normalise by system temperature only (tsys); $\frac{1}{T_{sys}^2}$ - this would be the case when $d\nu$, dt are known to be invariant
3. normalise by integration time only (tint): dT ; when T_{sys} and $d\nu$ are known to be invariant
4. normalise by variance - an empirical calculation from the spectra, that should implicitly capture any variations of any of the terms.

Task `tsdsmooth` applies a standard smoothing in the spectral domain, using (for now) a Gaussian kernel. Note the smoothing is done by convolution in the Fourier domain and therefore masking and blanking data can affect the outcome. Please use smooth masked data with care. For best results. baselining the data before smoothing to ensure there are few large excursions to zero in the spectral domain, which might contaminate the transformed data.

The task `sdmath` can perform basic arithmetic operations on *Scantable* or measurement set format datasets. This task will be unsupported in the near future. The `expr` paramter is formed using combinations of operators and optionally, datasets or variables; For two datasets containing, for example, "on" data and "off" data (i.e. science and reference observations), the data can be calibrated fvia the standard "on-off/off" algorithm:

```
expr='("orion_on_data.asap"-"orion_off_data.asap")/"orion_off_data.asap"
outfile='orion_cal.asap'
```

Task `sdscale` simply scales the data by a user-supplied constant value. This function is actually redundant with `sdmath` and also with `gencal` and `applycal`, and like `sdmath`, it will be unsupported in the near future.

8.5.3 Fitting: `sdfit`, `tsdfit`

From the user point of view, one of the key differences between `sdfit` and `tsdfit` is that `sdfit` applies additional smoothing, and is not part of the transition to measurement set (i.e. it can operate on *Scantable* format and Measurement set format). In the near future, `sdfit` will be unsupported.

In task `tsdfit`, fitting of (for now) gaussian profiles only, is done on a per-spectra basis, using iterative calls to a "line-finding" algorithm that passes the first guess for the requested number of gaussians, into the line fitting algorithm. Both steps are done in the image domain.

8.5.4 Dataset output: `sdplot`, `sdlist`, `sdstat`

These three tasks are available for quick, or more detailed analysis of single-dish data are largely superseded by `plotms`, `listobs` and `imstat`, which can do much of the same functionality with some exceptions.

`sdlist`, used to obtain basic observational information about the dataset is superseded by `listobs` for measurement set format data, but otherwise completes the same functionality. `sdstat` used to obtain statistical information about the measurement set data has additional capability for averaging that is lacking in `imstat`. Similarly, `sdplot` which is used to plot the spectral information of the data set has expanded capability for on-the-fly averaging before plotting that is lacking in `plotms`.

These tasks will be maintained until the migration to measurement set format is complete

8.5.5 Data input and output: `sdsave`

Data can be transported across formats using `sdsave`. Output formats include *Scantable*, Measurement set, SDFITS and ASCII. This task has little, if any application in core reduction of single-dish data, but is useful if the user wishes to export data into ASCII format for offline processing, outside CASA.

8.6 Import of NRO data

Importing NRO data is available, and both NEWSTAR and NOSTAR formats are supported, although multibeam data is currently problematic, and older data using AOS may show some inconsistencies relative to the original (before-import) data. CASA currently supports Dual-polarization data.

Note that an individual IFNO is assigned to each of arrays to identify data from an array. Note that arrays are numbered by successive IFNOs. For example, if observation of three arrays, A01, A03, and A05, are stored in a data set, their IFNOs will be 0, 1, and 2, respectively.

The parameter, `freqref`, controls how frequency reference frame is set. The default is 'rest'. The other option is 'vref'. If `freqref` is 'vref', frequency reference frame takes from VREF field in input NRO data. This parameter is only available in tool level, i.e. `sd.scantable`.

Note that at this point, processing of full polarization data needs to be done very carefully, and with particular attention paid to the spectral window identification, which is unreliably transported from NRO format to Measurement set format.

8.7 Using The ASAP Toolkit

ASAP is included with the CASA installation/build. It is loaded upon start-up, and the ASAP functionality is under the Python 'sd' tool.

The ASAP interface is essentially the same as that of the CASA toolkit, that is, there are groups of functionality (aka tools) which have the ability to operate on your data. The complete list of 'sd' tools are:

<code>sd.AsapLog</code>	<code>sd.asaplinefind</code>	<code>sd.mask_or</code>
<code>sd.__builtins__</code>	<code>sd.asaplog</code>	<code>sd.matplotlib</code>
<code>sd.__class__</code>	<code>sd.asaplog_post_dec</code>	<code>sd.merge</code>
<code>sd.__date__</code>	<code>sd.asapmath</code>	<code>sd.new_asaplot</code>
<code>sd.__delattr__</code>	<code>sd.asapplotter</code>	<code>sd.opacity</code>
<code>sd.__dict__</code>	<code>sd.average_time</code>	<code>sd.opacity_model</code>
<code>sd.__doc__</code>	<code>sd.calfs</code>	<code>sd.os</code>
<code>sd.__file__</code>	<code>sd.calibrate</code>	<code>sd.page</code>
<code>sd.__format__</code>	<code>sd.calnod</code>	<code>sd.parameters</code>
<code>sd.__getattr__</code>	<code>sd.calps</code>	<code>sd.plotter</code>
<code>sd.__hash__</code>	<code>sd.commands</code>	<code>sd.plotter2</code>
<code>sd.__init__</code>	<code>sd.coordinate</code>	<code>sd.pylab</code>
<code>sd.__name__</code>	<code>sd.dosigref</code>	<code>sd.quotient</code>
<code>sd.__new__</code>	<code>sd.dototalpower</code>	<code>sd.rc</code>
<code>sd.__package__</code>	<code>sd.edgemarker</code>	<code>sd.rcParams</code>
<code>sd.__path__</code>	<code>sd.env</code>	<code>sd.rcParamsDefault</code>
<code>sd.__reduce__</code>	<code>sd.fitter</code>	<code>sd.rcp</code>
<code>sd.__reduce_ex__</code>	<code>sd.flagplotter</code>	<code>sd.re</code>
<code>sd.__repr__</code>	<code>sd.get_revision</code>	<code>sd.sbseparator</code>
<code>sd.__revision__</code>	<code>sd.gui</code>	<code>sd.scantable</code>
<code>sd.__setattr__</code>	<code>sd.inspect</code>	<code>sd.selector</code>
<code>sd.__sizeof__</code>	<code>sd.interactivemask</code>	<code>sd.setup_env</code>
<code>sd.__str__</code>	<code>sd.ipysupport</code>	<code>sd.simplelinefinder</code>
<code>sd.__subclasshook__</code>	<code>sd.is_asap_cli</code>	<code>sd.skydip</code>
<code>sd.__version__</code>	<code>sd.is_casapy</code>	<code>sd.splitant</code>
<code>sd._asap</code>	<code>sd.is_ipython</code>	<code>sd.srctype</code>
<code>sd._is_sequence_or_number</code>	<code>sd.linecatalog</code>	<code>sd.sys</code>
<code>sd._n_bools</code>	<code>sd.linefinder</code>	<code>sd.toggle_verbose</code>
<code>sd._to_list</code>	<code>sd.list_files</code>	<code>sd.unique</code>
<code>sd.almacal</code>	<code>sd.list_rcparameters</code>	<code>sd.utils</code>
<code>sd.apexcal</code>	<code>sd.list_scans</code>	<code>sd.version</code>

```
sd.asapfitter          sd.logging           sd.welcome
sd.asapgrid           sd.mask_and         sd.xyplotter
sd.asapgrid2          sd.mask_not
```

A contextual description of the use and operation of these are given in Section 8.7.1

8.7.1 ASAP utility and tool descriptions

8.7.1.1 Rasterutil

Rasterutil is a module which enables you to select individual raster rows or rasters from raster-scanned dataset. Suppose you have a *Scantable* 'foo.asap' in which several rasters are scanned. To know how many rasters or raster rows this dataset contains, execute first as follows:

```
s=sd.scantable('foo.asap', False)
import rasterutil
r=rasterutil.Raster(s)
r.detect() .
```

Now you are ready to go ahead. Numbers of raster rows or rasters are shown by typing 'r.nrow' or 'r.nraster', respectively. Once you run `rasterutil.Raster.detect()`, IDs are given in chronological order both for each raster rows and rasters: a raster row can be specified by ID in range from 0 to `r.nrow-1`, while a raster can be specified by ID from 0 to `r.nraster-1`.

Here are sample commands to obtain a *Scantable* that contains specified raster row or raster only or other stuffs to select the specified ones:

```
srow0=r.asscantable(rowid=0)    #get \emph{Scantable} object containing the first raster row
sras1=r.asscantable(rasterid=1) #get \emph{Scantable} object containing the second raster
selrow2=r.aselector(rowid=2)   #get selector object for selecting the third raster row
taqlras3=r.astaql(rasterid=3)  #get TaQL query for selecting the 4th raster
```

Other useful commands include `rasterutil.Raster.plot_rows()`; data selected as a raster row or raster can be visualized using `asapplotter`.

Within ASAP, data is stored in a *Scantable*, which holds all of the observational information and provides functionality to manipulate the data and information. The building block of a *Scantable* is an integration which is a single row of a *Scantable*. Each row contains just one spectrum of a beam, IF and polarization.

Once you have a *Scantable* in ASAP, you can select a subset of the data based on scan numbers, or source names; note that each of these selections returns a new 'scantable' with all of the underlying functionality:

```
CASA <5>: scan27=scans.get_scan(27)          # Get the 27th scan
CASA <6>: scans20to24=scans.get_scan(range(20,25)) # Get scans 20 - 24
CASA <7>: scansOrion=scans.get_scan('Ori*')    # Get all Orion scans
```

To copy a scantable, do:

```
CASA <15>: ss=scans.copy()
```

8.7.1.2 Data Selection

In addition to the basic data selection above, data can be selected based on IF, beam, polarization, scan number as well as values such as T_{sys} . To make a selection create a `selector` object choose among various selection functions, e.g.,

```
sel = sd.selector()      # initialize a selector object
                        # sel.<TAB> will list all options
sel.set_ifs(0)          # select only the first IF of the data
scans.set_selection(sel) # apply the selection to the data
print scans             # shows just the first IF
```

8.7.1.3 State Information

Some properties of a *Scantable* apply to all of the data, such as spectral units, frequency frame, or Doppler type. This information can be set using the `scantable.set_xxxx` methods. These are currently:

```
CASA <1>: sd.scantable.set_<TAB>
sd.scantable.set_dirframe      sd.scantable.set_selection
sd.scantable.set_doppler      sd.scantable.set_sourcename
sd.scantable.set_feedtype     sd.scantable.set_sourcetype
sd.scantable.set_fluxunit     sd.scantable.set_spectrum
sd.scantable.set_freqframe    sd.scantable.set_T$_{\rm sys}$
sd.scantable.set_instrument   sd.scantable.set_unit
sd.scantable.set_restfreqs
```

For example, `sd.scantable.set_fluxunit` sets the default units that describe the flux axis:

```
scans.set_fluxunit('K') # Set the flux unit for data to Kelvin
```

Choices are 'K' or 'Jy'. Note: the `scantable.set_fluxunit` function only changes the **name** of the current fluxunit. To change fluxunit, use `scantable.convert_flux` as described in § 8.7.2.2 instead (currently it is necessary to do some gymnastics for non-AT telescopes).

Use `sd.scantable.set_unit` to set the units to be used on the spectral axis:

```
scans.set_unit('GHz') # Use GHz as the spectral axis for plots
```

The choices for the units are 'km/s', 'channel', or '*Hz' (e.g. 'GHz', 'MHz', 'kHz', 'Hz'). This does the proper conversion using the current frame and Doppler reference as can be seen when the spectrum is plotted.

Set the frame in which the frequency (spectral) axis is defined by `sd.scantable.set_freqframe`:

```
CASA <2>: help(sd.scantable.set_freqframe)
Help on method set_freqframe in module asap.scantable:
```

```
set_freqframe(self, frame=None) unbound asap.scantable.scantable method
  Set the frame type of the Spectral Axis.
  Parameters:
    frame:  an optional frame type, default 'LSRK'. Valid frames are:
            'REST', 'TOPO', 'LSRD', 'LSRK', 'BARY',
            'GEO', 'GALACTO', 'LGROUP', 'CMB'
  Examples:
    scan.set_freqframe('BARY')
```

The most useful choices here are `frame = 'LSRK'` and `frame = 'TOPO'` (what ALMA actually observes in). Note that the 'REST' option is not yet available. The Doppler frame is set with `sd.scantable.set_doppler`:

```
CASA <3>: help(sd.scantable.set_doppler)
Help on method set_doppler in module asap.scantable:
```

```
set_doppler(self, doppler='RADIO') unbound asap.scantable.scantable method
  Set the doppler for all following operations on this scantable.
  Parameters:
    doppler:  One of 'RADIO', 'OPTICAL', 'Z', 'BETA', 'GAMMA'
```

Finally, there are a number of functions to query the state of the scantable. These can be found in the usual way:

```
CASA <4>: sd.scantable.get_<TAB>
sd.scantable.get_abcissa      sd.scantable.get_parangle
sd.scantable.get_antennaname  sd.scantable.get_restfreqs
sd.scantable.get_azimuth     sd.scantable.get_rms
sd.scantable.get_column_names sd.scantable.get_row
sd.scantable.get_coordinate   sd.scantable.get_row_selector
sd.scantable.get_direction    sd.scantable.get_scan
sd.scantable.get_directionval sd.scantable.get_selection
sd.scantable.get_elevation    sd.scantable.get_sourcename
sd.scantable.get_fit          sd.scantable.get_spectrum
sd.scantable.get_fluxunit     sd.scantable.get_time
sd.scantable.get_inttime      sd.scantable.get_T${\rm sys}$
sd.scantable.get_mask         sd.scantable.get_T${\rm sys}$spectrum
sd.scantable.get_mask_indices sd.scantable.get_unit
sd.scantable.get_masklist     sd.scantable.get_weather
```

These include functions to get the current values of the states mentioned above, as well as methods to query the number of scans, IFs, and polarizations in the *Scantable* and their designations. See the inline help of the individual functions for more information.

8.7.1.4 Masks

Several functions (fitting, baseline subtraction, statistics, etc.) may be run on a range of channels (or velocity/frequency ranges). You can create masks of this type using the `create_mask` function:

```
# spave = an averaged spectrum
spave.set_unit('channel')
rmsmask=spave.create_mask([5000,7000]) # create a region over channels 5000-7000
rms=spave.stats(stat='rms',mask=rmsmask) # get rms of line free region

rmsmask=spave.create_mask([3000,4000],invert=True) # choose the region
# *excluding* the specified channels
```

The mask is stored in a simple Python variable (a list) and so can be manipulated using Python facilities.

8.7.1.5 *Scantable* Management

Scantable can be listed via:

```
CASA <33>: sd.list_scans()
The user created scantables are:
['scans20to24', 's', 'scan27']
```

As every *Scantable* will consume memory usage, if you will not use it any longer, you can explicitly remove it via:

```
del <scantable name>
```

8.7.1.6 *Scantable* Mathematics

It is possible to do simple mathematics directly on scantables from the CASA command line using the `+`, `-`, `*`, `/` operators as well as `+=`, `-=`, `*=`, `/=`

```
CASA <10>: scan2=scan1+2.0 # add 2.0 to data
CASA <11>: scan *= 1.05    # scale spectrum by 1.05
```

Operands can be a numerical value and one- or two-dimensional Python list. For list operand, its shape should be conform with the shape of spectral data stored in the scantable. Mathematics between two scantables is also available. In that case, scantables must be conform with each other.

NOTE: In scantable mathematics, scantable must be put on the left. For example:

```
CASA<12>: scan2=scan1+2.0    # this works
CASA<13>: scan2=2.0+scan1    # this causes an error
```


8.7.1.7 *Scantable Save and Export*

ASAP can export scantables in a variety of formats, suitable for reading into other packages. The formats are:

- ASAP – This is the internal format used for ASAP. It is the only format that allows the user to restore the data, fits, etc., without losing any information. The ASAP *Scantable* is a CASA Table (memory-based table; but can be made as temporarily disk-based with `sd.rcParam['scantable.storage']`). This function just converts it to a disk-based table. You can access it with the CASA `browsetable` task or any other CASA table tasks.
- SDFITS – The Single-dish FITS format. This format was designed for interchange between packages but few packages can actually read it.
- ASCII – A simple text based format suitable for the user to process using Python or other means.
- Measurement Set (V2: CASA format) – Saves the data in a Measurement Set. All CASA tasks which use an Measurement Set should work with this format.

Scantables are exported by the function, `save`:

```
CASA : scans.save(name='foo.output.ms',format='MS2')
```

8.7.2 Calibration

For some observatories, the calibration happens transparently as the input data contains the T_{sys} measurements taken during the observations. The nominal ' T_{sys} ' values may be in Kelvin or Jansky. The user may wish to apply a T_{sys} correction or apply gain-elevation and opacity corrections.

8.7.2.1 T_{sys} scaling

If the nominal T_{sys} measurement at the telescope is wrong due to incorrect calibration, the `scale` function allows it to be corrected.

```
scans.scale(1.05,T${\rm sys}$=True) \# by default only the spectra are scaled
\# (and not the corresponding T${\rm sys}$) unless T${\rm sys}$=True
```

8.7.2.2 Flux and Temperature Unit Conversion

The function, `convert_flux`, is available for converting measurements in Kelvin to Jansky (and vice versa). It converts and scales data to the selected units. The user may need to supply the aperture efficiency, telescope diameter, or the Jy/K factor

```
scans.convert_flux(eta=0.48, d=35.) # Unknown telescope
scans.convert_flux(jypk=15) # Unknown telescope (alternative)
scans.convert_flux() # known telescope (mostly AT telescopes)
scans.convert_flux(eta=0.48) # if telescope diameter known
```

Chapter 9

Simulation

New in 4.4:

- Single dish imaging is accomplished in `simalma` and `simobserve` by calling the `sdimaging` task with standard imaging parameters in ALMA, i.e., with spheroidal gridding, the cell size of 1/9 of the FWHM of primary beam, and convolution support of 6 pixes.

The tasks available for simulating observations are:

- `simobserve` — simulate an interferometer or total power observation (§ 9.1)
- `simanalyze` — image and analyze simulated data sets (§ 9.1)
- `simalma` — simulate an ALMA observation including multiple configurations of the 12-m interferometric array, the 7-m ACA, and total power measurements. Generate a combined image from the simulated data sets (§ 9.2)

The capability of simulating observations and data sets from the JVL A and ALMA are an important use-case for CASA. This not only allows one to get an idea of the capabilities of these instruments for doing science, but also provides benchmarks for the performance and utility of the software to process “realistic” data sets (with atmospheric and instrumental effects). Simulations can also be used to tune parameters of the data reduction and therefore help to optimize the process. CASA can calculate visibilities (create a measurement set) for any interferometric array, and calculate and apply calibration tables representing some of the most important corrupting effects. `simobserve` can also simulate total power observations, which can be combined with interferometric data in `simanalyze` (i.e. one would run `simobserve` twice, `simanalyze` once). The task `simalma` is a task to simulate an ALMA observation, including ALMA 12-m, ACA 7-m and total power arrays, and generate a combined image. `simalma` also attempts to provide useful feedback on those different observation components, to help the user better understand the observing considerations.

CASA’s simulation capabilities continue to be improved with each CASA release. For the most current information, please refer to <http://www.casaguides.nrao.edu>,

Inside the Toolkit:

The simulator methods are in the `sm` tool. Many of the other tools are also helpful when constructing and analyzing simulations.

and click on “Simulating Observations in CASA”. Following general CASA practice, the greatest flexibility and richest functionality is at the Toolkit level. The most commonly used procedures for interferometric and single dish simulation are encapsulated in the `simobserve` task.

9.1 Simulating ALMA observations with `simobserve` and `simanalyze`

The `simobserve` inputs are (submenus expand slightly differently for `thermalnoise=manual` and single dish observing):

```

project          =      'sim'          # root prefix for output file names
skymodel         =      ''            # model image to observe
  inbright       =      ''            # scale surface brightness of
                                     # brightest pixel e.g. "1.2Jy/pixel"
  indirection    =      ''            # set new direction
                                     # e.g. "J2000 19h00m00 -40d00m00"
  incell         =      ''            # set new cell/pixel size
                                     # e.g. "0.1arcsec"
  incenter       =      ''            # set new frequency of center
                                     # channel e.g. "89GHz"
                                     # (required even for 2D model)
  inwidth        =      ''            # set new channel width
                                     # e.g. "10MHz" (required even
                                     # for 2D model)

complist         =      ''            # componentlist to observe
  compwidth      =      '8GHz'        # bandwidth of components

setpointings     =      True          #
  integration    =      '10s'        # integration (sampling) time
  direction      =      ''            # "J2000 19h00m00 -40d00m00"
                                     # or "" to center on model
  mapsize        =      ['', '']     # angular size of map or ""
                                     # to cover model
  maptype        =      'ALMA'        # hexagonal, square (raster),
                                     # ALMA, etc.
  pointingspacing =      ''          # spacing in between
                                     # pointings or "0.25PB" or ""
                                     # for ALMA default
                                     # INT=lambda/D/sqrt(3), SD=lambda/D/3

obsmode          =      'int'         # observation mode to simulate
                                     # [int(interferometer)|sd(singledish)|""(none)]
  antennalist    =      'alma.out10.cfg' # interferometer antenna position file
  refdate        =      '2014/05/21' # date of observation - not
                                     # critical unless concatenating
                                     # simulations

```

```

hourangle      = 'transit'      # hour angle of observation
                                     # center e.g. "-3:00:00", "5h", "-4.5" (a
                                     # number without units will
totaltime      = '7200s'        # total time of observation
                                     # or number of repetitions
caldirection   = ''            # pt source calibrator [experimental]
calflux        = '1Jy'
thermalnoise   = 'tsys-atm'     # add thermal noise:
                                     # [tsys-atm|tsys-manual|""]
user_pwv       = 1.0            # Precipitable Water Vapor in mm
t_ground       = 269.0          # ambient temperature
seed           = 11111          # random number seed

leakage        = 0.0            # cross polarization (interferometer only)
graphics       = 'both'         # display graphics at each
                                     # stage to [screen|file|both|none]

verbose        = False
overwrite      = True           # overwrite files starting with $project

```

This task takes an input model image or list of components, plus a list of antennas (locations and sizes), and simulates a particular observation (specified by mosaic setup and observing cycles and times). The output is a measurement set suitable for further analysis in CASA.

The `simanalyze` inputs are:

```

project        = 'sim'          # root prefix for output file names
image          = True           # (re)image $project.*.ms to $project.image
vis            = 'default'      # Measurement Set(s) to image
modelimage     = ''            # lower resolution prior
                                     # image to use in clean e.g. existing total
                                     # power image
imsize         = 0              # output image size in pixels
                                     # (x,y) or 0 to match model
imdirection    = ''            # set output image direction,
                                     # (otherwise center on the model)
cell           = ''            # cell size with units or "" to equal model
interactive    = False         # interactive clean? (make
                                     # sure to set niter>0 also)
niter          = 0              # maximum number of
                                     # iterations (0 for dirty image)
threshold      = '0.1mJy'      # flux level (+units) to stop cleaning
weighting      = 'natural'      # weighting to apply to
                                     # visibilities briggs will use robust=0.5
mask           = []            # Cleanbox(es), mask
                                     # image(s), region(s), or a level
outertaper     = []            # uv-taper on outer baselines in uv-plane
stokes         = 'I'           # Stokes params to image
featherimage   = ''            # image (e.g. total power) to

```

```

# feather with new image

analyze          =      True      # (only first 6 selected
# outputs will be displayed)
  showuv         =      True      # display uv coverage
  showpsf       =      True      # display synthesized (dirty)
# beam (ignored in single dish simulation)
  showmodel      =      True      # display sky model at original resolution
  showconvolved =      False     # display sky model convolved with output beam
  showclean     =      True      # display the synthesized image
  showresidual  =      False     # display the clean residual
# image (ignored in single dish simulation)
  showdifference =      True      # display difference between
# output cleaned image and input
# model sky image convolved
# with output clean beam
  showfidelity  =      True      # display fidelity

graphics        =      'both'    # display graphics at each
# stage to [screen|file|both|none]

verbose         =      False
overwrite       =      True      # overwrite files starting with $project

```

This task analyzes one or more measurement sets - interferometric and/or single dish. The output is a synthesized image created from those visibilities, a difference image between the synthesized image and your sky model convolved with the output synthesized beam, and a fidelity image. (see ALMA memo 398 for description of fidelity, which is approximately the output image divided by the difference between input and output)

9.2 Simulating ALMA observations with `simalma`

The task `simalma` simulates an ALMA observation by ALMA 12-m, ACA-7m and total power arrays. It takes an input model image or a list of components, plus configurations of ALMA antennas (locations and sizes), and simulates a particular ALMA observation (specified by mosaic setup and observing cycles and times). The outputs are measurement sets. The task optionally generates synthesized images from the measurement sets as `simanalyze` does.

Technically speaking, `simalma` internally calls `simobserve` and `simanalyze` as many times as necessary to simulate and analyze an ALMA observation. Some of the simulation (`simobserve`) and imaging (`simanalyze`) parameters are automatically set to values typical of ALMA observations in `simalma` (see § 9.2.1 for more details). Thus, it has a simpler task interface compared to `simobserve` plus `simanalyze` at the cost of limited flexibility. If you want to have more control on simulation setup, it is available by manually running `simobserve` and `simanalyze` multiple times or by using `sm` tools.

The `simalma` inputs are:

```

project          =      ''          # root prefix for output file names
project          =      'sim'       # root prefix for output file names
dryrun           =      False      # dryrun=True will only
                                     # produce the informative report, not run
                                     # simobserve/analyze
skymodel         =      ''          # model image to observe
  inbright       =      ''          # scale surface brightness of
                                     # brightest pixel e.g. "1.2Jy/pixel"
  indirection    =      ''          # set new direction
                                     # e.g. "J2000 19h00m00 -40d00m00"
  incell         =      ''          # set new cell/pixel size e.g. "0.1arcsec"
  incenter       =      ''          # set new frequency of center
                                     # channel e.g. "89GHz" (required even
                                     # for 2D model)
  inwidth        =      ''          # set new channel width
                                     # e.g. "10MHz" (required even for 2D model)

complist         =      ''          # componentlist to observe
  compwidth      =      '8GHz'      # bandwidth of components

setpointings     =      True        #
  integration    =      '10s'       # integration (sampling) time
  direction      =      ''          # "J2000 19h00m00 -40d00m00"
                                     # or "" to center on model
  mapsize        =      ['', '']    # angular size of map or "" to cover model

graphics         =      'both'      # display graphics at each
                                     # stage to [screen|file|both|none]

verbose          =      False       #
overwrite        =      False       # overwrite files starting with $project

```

The task `simalma` is designed as a task that is invoked only once for a simulation setup. It always sets up `skymodel` and `pointings`. That means that `simalma` is not supposed to be run multiple times for a project, unlike `simobserve` and `simanalyze`. The task `simalma` may ignore or overwrite the old results when it is run more than once with the same project name.

There are options in `simalma` to simulate observation of ACA 7-m and total power arrays, to apply thermal noise, and/or to generate images from simulated measurement sets. One inputs a vector of configurations, and a corresponding vector of `totaltimes` to observe each component. Thermal noise is added to visibilities when `pwv > 0`. The ATM atmospheric model is constructed from the characteristics of the ALMA site and a user defined Precipitable Water Vapour (`pwv`) value. Set `pwv = 0` to omit the thermal noise. Finally, when `image = True`, synthesized images are generated from the simulated measurement sets.

9.2.1 Implementation details

As mentioned in the previous section, `simalma` automatically sets some of the simulation and imaging parameters to values typical of ALMA observations. The implementations of antenna

configurations, pointings, integration time, and imaging in CASA 4.4.0 are described in this section.

Antenna Configuration:

The configurations of the ALMA 12-m and 7-m arrays are defined by the `antennalist` parameter, which can be a vector. Each element of the vector can be either the name of an antenna configuration file or a desired resolution, e.g., `'alma;cycle1;5arcsec'`. Some examples:

- `antennalist = ['alma.cycle2.5.cfg', 'aca.cycle2.i.cfg']; totaltime = ['20min', '2h']`: Will observe the 12-m array in configuration C32-5 for 20 minutes and the ACA 7-m array for 2 hours.
- `antennalist = ['alma;cycle2;0.5arcsec', 'aca.i.cfg']; totaltime = ['20min', '2h']`: Will observe the 12-m array in whatever cycle 2 configuration yields a zenith synthesized beam as close as possible to 0.5 arcsec (at the center frequency of your skymodel) for 20 minutes and the ACA 7-m array for 2 hours.
- `antennalist = ['alma.cycle1.2.cfg', 'aca.cycle2.i.cfg']; totaltime = '20min'`: Will observe the 12-m array in cycle 1 configuration for 20 minutes and the ACA 7-m array for the default of $2 \times (12\text{-m time}) = 1\text{h}20\text{min}$. This parameter setting will also generate a warning that the user is combining configurations from different ALMA Cycles (but the simulation will run despite that).

Total power can either be included along with interferometric configurations e.g. `antennalist = ['alma.cycle1.2.cfg', 'aca.cycle2.i.cfg', 'alma.tp.cfg']`, or by using the `tpnant` and `tptime` parameters. The latter is preferred since it allows greater control (in particular the number of total power antennas to use – if more than one is used, multiple total power observations will be generated and combined in imaging).

Field Setup:

There are two ways to setup pointings, i.e., Rectangle Setup and Multi-Pointing.

In the Rectangle Setups, pointings are automatically calculated from the pointing centre (`direction`) and the map size. A rectangular map region is covered by a hexagonal grid (`maptype = 'alma'`) with Nyquist sampling, i.e., 0.48 PB spacing (where $\text{PB} \equiv 1.2 \lambda/D$), in both ALMA 12-m and ACA 7-m array simulations. A slightly larger area is mapped in ACA total power simulations for later combination with interferometer visibilities. The map area is extended by 1 PB in each direction and covered by a lattice grid with 0.225 PB spacing.

In Multi-Pointing, a list of pointings is defined in the `direction` parameter or read from a file (when `setpointings = False`). The ALMA 12-m and ACA 7-m arrays observe the specified directions. The ACA total power simulations map either (1) square regions of 2 PB extent centred at each of the pointings, or (2) a rectangle region that covers all the pointings. Either (1) or (2), whichever can be done with the smaller number of points, is selected. The pointing spacing in total power simulations is, again, 0.225 PB in lattice grids.

It is advisable that for Total Power Simulations, the field is chosen sufficiently large, maybe padding at least 1-2 primary beams on each side.

Integration time:

The total observation time of each component or configuration is defined by the `totaltime` parameter as noted above. A scalar will trigger use of the Cycle 2 default time multipliers, 1:0.5:2:4 for the first 12-m configuration, any additional 12-m configurations, any 7-m configuration, and any total power observation.

In general, the integration time (dump interval) of simulations is defined by the `integration` parameter with an exception. Since the ACA total power array always observes larger areas compared to the ALMA 12-m and ACA 7-m arrays, it is possible that the ACA total power array cannot cover all pointings in the given observation time. In such a case, the integration time in the total power simulation is scaled so that the all pointings are observed at least once in its observation time, i.e., $\text{integration_TP} = \text{tptime} / (\text{the number of total power pointings})$.

Imaging and combination of ALMA with ACA:

The CLEAN algorithm is used in `simalma` to generate images from visibilities. The visibilities are weighted to UV-plane using Briggs weighting.

When ACA observations are simulated, visibilities of ACA 7-m are weighted by the relative sensitivities to ALMA 12-m visibilities, and both data sets are concatenated before imaging. The relative weight of ACA 7-m visibilities is defined in proportion to the difference of beam area, i.e., $(7/12)^2 = 0.34$. This is because `simalma` uses a bandwidth and an integration time common to both ALMA 12-m and ACA 7-m simulations.

The interferometer and total power images are combined using `feather` task when total power observations are included. The total power image is scaled by the interferometer primary beam coverage before combination. The final image product is the combined image corrected for the interferometer primary beam coverage. The output image of the `feather` task is divided by the interferometer primary beam coverage in the final step.

Chapter 10

Parallel Processing in CASA

Starting in CASA 4.5.0, a *parallelized* execution of a full data analysis from data import to imaging is possible using a new infrastructure that is based on the Message Passing Interface (MPI). Briefly, MPI is a standard which addresses primarily the message-passing parallel programming model in a practical, portable, efficient and flexible way.

10.1 The CASA parallelization scheme

In order to run *one* analysis on multiple processors, one can parallelize the work by dividing the data into several parts (“partitioning”) and then run a CASA instance on each part *or* have non-trivially parallelized algorithms, which make use of several processors within a single CASA instance. Non-trivial parallelization is presently only implemented in certain sections of the imaging code of CASA based on OpenMP.

All other parallelization is achieved by partitioning the MeasurementSet (MS) of interest using the task `partition` or at import time using `importasdm`. The resulting partitioned MS is called a “Multi-MS” or “MMS”. Logically, an MMS has the same structure as an MS but internally it is a group of several MSs which are virtually concatenated. Virtual concatenation of several MSs or MMSs into an MMS can also be achieved via task `virtualconcat`.

Due to the virtual concatenation, the main table of an MMS appears like the union of the main tables of all the member MSs such that when the MMS is accessed like a normal MS, processing can proceed sequentially as usual. Each member MS or “Sub-MS” of an MMS, however, is at the same time a valid MS on its own and can be processed as such. This is what happens when the MMS is accessed by a parallelized task. The partitioning of the MMS is recognized and work is started in parallel on the separate Sub-MSs, provided that the user has started CASA with `mpicasa`. See how to start CASA in parallel in § 10.3.1.

The internal structure of an MMS can be inspected using the task `listpartition`. See § 2.2.8.

10.2 Multi-MS creation

10.2.1 Partition

The `partition` task is the main task to create a “Multi-MS”. It takes an input Measurement Set and creates an output “Multi-MS” based on the data selection parameters.

The inputs to `partition` are:

```
# partition :: Task to produce Multi-MSs using parallelism
vis                =      ''          # Name of input measurement set
outputvis          =      ''          # Name of output measurement set
createmms         =      True         # Should this create a multi-MS output
  separationaxis   =      'auto'       # Axis to do parallelization across(scan, spw, auto)
  numsubms        =      'auto'       # The number of SubMSs to create (auto or any number)
  flagbackup      =      True         # Create a backup of the FLAG column in the MMS.

datacolumn        =      'all'        # Which data column(s) to process.
field             =      ''           # Select field using ID(s) or name(s).
spw               =      ''           # Select spectral window/channels.
scan              =      ''           # Select data by scan numbers.
antenna           =      ''           # Select data based on antenna/baseline.
correlation       =      ''           # Correlation: '' ==> all, correlation='XX,YY'.
timerange         =      ''           # Select data by time range.
intent            =      ''           # Select data by scan intent.
array             =      ''           # Select (sub)array(s) by array ID number.
uvrange           =      ''           # Select data by baseline length.
observation        =      ''           # Select by observation ID(s).
feed              =      ''           # Multi-feed numbers: Not yet implemented.
```

10.2.1.1 The `createmms` parameter

The keyword `createmms` is by default set to `True` to create an output MMS. It contains three sub-parameters, `separationaxis`, `numsubms` and `flagbackup`. `Partition` accepts three axis to do separation across: `'auto'`, `'scan'` or `'spw'`. The default `separationaxis='auto'` will first separate the MS in spws, then in scans. It tries to balance the spw and scan content in each Sub-MS also taking into account the available fields.

The user may force the number of “Sub-MSs” in the output MMS by setting the sub-parameter `numsubms`. The default `'auto'` is to create as many “Sub-MSs” as the number of engines used when starting CASA, in an optimized way. Details on how to set the number of engines using `mpicasa` are given in §10.3.2.

The `flagbackup` sub-parameter will create a backup of the FLAG column and save it to the `.flagversions` file.

10.2.2 Importasdm

Task `partition` has been embedded in task `importasdm` so that at import time the user can already create a MMS. Set the parameter `createmms` to `True` and the output of `importasdm` will be a MMS created with default parameters. Sub-parameters `separationaxis` and `numsubms` are also available in `importasdm`. From this point on in the data reduction chain, tasks that have been parallelized will run automatically in parallel when they see an MMS and tasks that are not parallelized will work in the same way as they normally do on a MS.

10.3 Parallelization control

10.3.1 Requirements

The following requirements are necessary for all the nodes to be included in the cluster:

- Password-less ssh access from the client (user) machine into all the hosts to be included in the cluster.

NOTE: This is not necessary when using only `localhost`, i.e. if the cluster is deployed only on the machine where `casapy` is running.

- All the input files must be located in a shared file-system, accessible from all the nodes comprising the cluster, and mounted in the same path of the file-system.
- Mirrored CASA installation w.r.t. the CASA installation in the client (user) machine, so that the following environmental variables are pointing to valid installations: `PATH`, `LD_LIBRARY_PATH`, `IPYTHONDIR`, `CASAPATH`, `CASAARCH`, `PYTHONHOME`, `__CASAPY_PYTHONDIR`, `PGPLOT_DEV`, `PGPLOT_DIR`, `PGPLOT_FONT`.

10.3.2 Configuration and Start-Up

The main library used in CASA (4.4+) to achieve parallelization is the Message Passing Interface (MPI). MPI is already included in the CASA distribution so that users do not need to install it. The CASA distribution comes with a wrapper of the MPI executor, which is called `mpicasa`. This wrapper does several settings behind the scenes in order to properly configure the environment to run CASA in parallel.

The “cluster”, i.e. the collection of CASA instances which will run the jobs from parallelized tasks, is set up via `mpicasa`. `Mpicasa` only needs to know how many nodes will be used in the processing. The simplest example is to run CASA in parallel on the `localhost` using the available engines in the machine. A typical example would be a desktop with 16 engines. If you want to use half of the engines to run CASA in parallel, give this information to `mpicasa` and call CASA normally. Example:

```
mpicasa -n <number_of_engines> path_to_casa/casa <casa_options>
```

Where:

1. `mpicasa`: Wrapper around `mpirun`, which can be found in the `casa` installation directory.
Example: `/home/user/casa-release-4.5.0-el6/`
2. `-n`: MPI option to get the number of engines to use in the processing.
3. `number_of_engines`: The number of engines to be used in the localhost machine.
NOTE: MPI uses one engine as the MPI Client, which is where the user will see messages printed in the terminal or in the logger. The other engines are used for the parallel work and are called MPI Servers. Because of this, usually we give `number_of_engines + 1`.
4. `casa`: Full path to the CASA executable, `casa` or `casapy`.
5. `casa_options`: CASA options such as: `-c`, `-nogui`, `-log2term`, etc.

It is also possible to use other nodes, which can form a “cluster”. Following the requirements from §10.3.1, replace the “`-n`” option of `mpicasa` with a “`-hostfile host_file`”, as shown below:

```
mpicasa -hostfile <host_file> path_to_casa/casa <casa_options>
```

Where:

1. `host_file`: It is a text file containing the name of the nodes forming the cluster and the number of engines to use in each one of the nodes.

Example:

```
orion slots=5
antares slots=4
sirius slots=4
```

The above configuration file will set up a cluster comprised of three nodes (orion, antares and sirius), deploying the engines per node as follows: At host “orion” up to 5 engines will be deployed (including the MPI Client). If the processing requires more engines, it will take them from “antares” and once all the 4 engines in “antares” are used, it will use up to 4 engines in “sirius”.

To get help do:

```
mpicasa --help
```

10.3.3 Examples of running CASA in parallel

The following is a list of typical examples on how to run CASA in parallel. Once CASA is started with `mpicasa` and the “Multi-MS” is created, there is no difference between running CASA in serial and in parallel. You can find an example of a parallelized analysis in the following regression script located in a sub-directory of your CASA distribution.

```
alma-m100-analysis-hpc-regression.py
```

Example 1. Run the above regression script in parallel, using 8 engines in parallel and 1 engine as the MPI Client.

```
mpicasa -n 9 path_to_casa/casa --nogui --log2term -c alma-m100-analysis-hpc-regression.py
```

Example 2. Start CASA as described in §10.3.2, for an interactive session, using 5 engines on the local machine.

```
mpicasa -n 5 path_to_casa/casa --nogui --log2term
```

Run `importasdm` to create a “Multi-MS” and save the online flags to a file. The output will be automatically named `uid__A002_X888a.ms`, which is an MMS partitioned across `spw` and `scan`. The online flags are saved in the file `uid__A002_X888a_cmd.txt`.

```
CASA <2>: importasdm('uid__A002_X888a', createmms=True, savecmds=True)
```

List the contents of the MMS using `listobs`. In order to see how the MMS is partitioned, use `listpartition`.

```
CASA <3>: listobs('uid__A002_X888a.ms', listfile='uid__A002_X888a.listobs')
CASA <4>: listpartition('uid__A002_X888a.ms')
```

Apply the online flags produced by `importasdm`, using `flagdata` in list mode. Flagdata is parallelized so that each engine will work on a separated “Sub-MS” to apply the flags from the `uid__A002_X888a_cmd.txt` file. You will see messages in the terminal (also saved in the `casapy-###.log` file), containing the strings `MPIServer-1`, `MPIServer-2`, etc., for all the engines that process in parallel.

```
CASA <5>: flagdata('uid__A002_X888a.ms', mode='list', infile='uid__A002_X888a_cmd.txt')
```

Flag auto-correlations and the high `Tsys` antenna also using list mode for optimization.

```
CASA <6>: flagdata('uid__A002_X888a.ms', mode='list',
                infile=["autocorr=True","antenna='DA62'"])
```

Create all calibration tables in the same way as for a normal MS. Task `gaincal` is not parallelized, therefore it will work on the MMS as if it was a normal MS.

```
CASA <7>: gaincal('uid__A002_X888a.ms', caltable='cal-delay_uid__A002_X888a.K',
                 field='*Phase*', spw='1,3,5,7', solint='inf', combine='scan',
                 refant=therefant, gaintable='cal-antpos_uid__A002_X888a',
                 gaintype='K')
```

Apply all the calibrations to the MMS. `applycal` will work in parallel on each “Sub-MS” using the available engines.

```
CASA <8>: applycal(vis='uid__A002_X888a.ms', field='0', spw='9,11,13,15',
                  gaintable=['uid__A002_X888a.tsys',
                             'uid__A002_X888a.wvr.smooth',
                             'uid__A002_X888a.antpos'],
                  gainfield=['0', '', ''], interp='linear,linear',
                  spwmap=[tsysmap, [], []], calwt=True, flagbackup=False)
```

Split out science spectral windows. The difference here is that we use `split2` instead of `split`. For the 4.5 release some of the parallelized tasks have an index 2 such as `split2`, `cvel2` and `hanningsmooth2`. Starting in version 4.6 these tasks will be renamed to `split`, `cvel` and `hanningsmooth`, respectively so that no other change will be needed when running a script in parallel. Tasks `split2`, `cvel2` and `hanningsmooth2` will recognize if the input is an MMS and will process it in parallel, creating also an output MMS.

```
CASA <9>: split2(vis='uid__A002_X888a.ms', outputvis='uid__A002_X888a.ms.split',
                datacolumn='corrected', spw='9,11,13,15', keepflags=True)
```

Run `clean` or `tclean` normally to create your images.

Appendix A

Obtaining, Installing, and Customizing CASA

A.1 Installation On Linux

To install CASA for Linux, we have packaged up a binary distribution of CASA which is available as a downloadable tar file. We believe this binary distribution works with most Linux distributions. While the binary distribution is the only supported public distribution, most CASA developers use RPMs for many third-party packages installed with yum to do development on RedHat Enterprise Linux. Installing the developer RPMs requires root access and we only provide developer support for organizations which have a cooperative agreement to participate in the development of CASA. We are currently working on the development of a distribution for developers similar to our standard binary distribution, but it is not yet ready for testing.

A.1.1 Installation

You do not have to have root or sudo permission, you can easily install CASA, delete it, move it, and it works for many versions of Linux. The one caveat is that CASA on Linux currently will not run if the Security-Enhanced Linux option of the linux operating system is set to enforcing. For the non-root install to work, SELinux must be set to disabled or permissive (in `/etc/selinux/config`) or you must run (as root):

```
setsebool -P allow_execheap=1
```

Otherwise, you will encounter errors like:

```
casapy: error while loading shared libraries:  
/opt/casa/casapy-20.0.5653-001/lib/liblapack.so.3.1.1:  
cannot restore segment prot after reloc: Permission denied
```


The non-root installation is thought to work on a wide variety of linux platforms, see Sect. 1.2 for the latest supported OSs.

A.1.1.1 Using more than one Linux version of CASA

Up to CASA 3.3.0, CASA .rpm files prohibited installing more than one CASA release at a time. Starting with CASA 3.4.0, CASA .rpm files allow previously installed CASA releases to remain installed.

To start a specific CASA version, type

```
casapy --release <VERSION>
```

or

```
casapy -r <VERSION>
```

where `<VERSION>` is a placeholder for the CASA version to be invoked, e.g. 3.3.0.

The programs `asdm2MS`, `casabrowser`, `casalogger`, `casaplotms`, `casapy`, `casapyinfo`, and `casaviewer` all take the two new command line options: `-r` and `--release`. These options allow users to select a CASA program to run from the installed CASA releases.

A.1.2 Unsupported platforms

The non-root install may work on other platforms not listed, please let us know if you find that this binary distribution of CASA works on other linux platforms. Also note, that the plotting tasks like `plotxy` and `plotcal` are the ones that typically give problems for new platforms, so a check of these after attempting an unsupported platform installation is advisable.

A.1.3 Download & Unpack

You can download the distribution tar file from

```
http://casa.nrao.edu/casa\_obtaining.shtml
```

This directory will contain two tar files one will be the 32-bit version of CASA and the other will be the 64-bit version of CASA. The file name of the 64-bit version ends with `-64b.tar.gz`. After downloading the appropriate tar file, `untar` it with

```
tar -zxf casapy-*.tar.gz
```

This will extract a directory with the same basename as the tar file. Change to that directory and add it to your path with, for example,

```
PATH='pwd':\${PATH}.
```

After that, you should be able to start CASA by running

```
casapy
```

A.2 Installation on Mac OS

CASA for Macintosh is distributed as self-contained Macintosh application. For installation purposes, this means that you can install CASA by simply dragging the application to your hard disk. It should be as easy as copying a file.

1. Download the CASA disk image for your OS version from our download site http://casa.nrao.edu/casa_obtaining.shtml
2. Open the disk image file (if your browser does not do so automatically).
3. Drag the CASA application to the Applications folder of your hard disk.
4. Eject the CASA disk image.
5. Double-click the CASA application to run it for the first time. This ensures everything is properly updated if you had installed a previous version.
6. Optional: Create symbolic links to the CASA version and its executables (Administrator privileges are required), which will allow you to run `casapy`, `casaviewer`, `casaplotms`, etc. from any terminal command line. To do so, run `!create-symlinks` from a CASA prompt. (In CASA 4.2.2 and earlier you will be prompted to update the symlinks on double clicking the Application)

You may need to unload the dbus before the copy will work

```
launchctl remove org.freedesktop.dbus-session
launchctl remove org.freedesktop.dbus-system
```

Versions after 12115 are 64bit only and will not work on older mac intel machines The first time you launch the CASA application, it will prompt you to set up an alias to the `casapy` command. You will be taken through the process of creating several `casapy` symbolic links, it is advisable to do so as this will allow you to run `casapy` from a terminal window by typing `casapy`. Additionally, the viewer (`casaviewer`), table browser (`casabrowser`), plotms (`casaplotms`), and `buildmytasks` will also be available via the command line. Creating the symbolic links will require that you have administrator privileges.

A.2.0.1 Using more than one Mac version of CASA

By dragging the CASA.app into the Applications folder, any previous version of CASA will be replaced. If one would like to keep older versions, one can simply rename them, e.g., to CASA-3.3.0.app. Double clicking any of the CASA*.app applications will prompt to update the symlinks to that specific CASA version. So any startup of `casapy`, `casaviewer`, `casaplotms` will point to that version. If one decides to switch to a different version, just double click the respective CASA*.app and follow the instructions to update the symlinks.

A.3 Startup

in a terminal type

```
casapy
```

and the world of CASA will open its doors for you.

There are a number of options to `casapy` (see `casapy --help`): Options are:

<code>--rcdir</code>	directory	
<code>--logfile</code>	logfile	specify the name of the log file if other than <code>casapy-DATE.log</code>
<code>--maclogger</code>		will use the Mac Console program for the logger
<code>--log2term</code>		output the logger text in the terminal
<code>--nologger</code>		run without launching the logger
<code>--nologfile</code>		does not create a logfile
<code>--nogui</code>		will not open the logger GUI
<code>--colors=[NoColor Linux LightBG]</code>		selects color theme for prompt task inputs
<code>--noipython</code>		does not launch ipython
<code>--release <VERSION></code>		launches CASA version <VERSION> when installed as Linux rpm
<code>-r <VERSION></code>		alias for <code>--release</code>
<code>-c filename-or-expression</code>		execute a CASA python script from the command line
<code>--help</code>		print this text and exit

E.g. you can execute a CASA script `script.py` directly with the command

```
casapy -c script.py
```

You can also launch the `plotms` and `viewer` GUIs separately without starting CASA itself. To do so, type:

```
casaplotms
```

to launch `plotms` and

```
casaviewer
```

for the viewer.

A.4 Startup Customization

There are two initialization files that are loaded upon startup. The first is loaded very early in the startup of `casapy`:

```
~/casa/prelude.py
```

This allows for limited customization of the `casapy` environment, e.g. setting the path to an alternate logger. The second startup file should be used for most purposes:

```
~/casa/init.py
```

This file is loaded just before the `casapy` prompt is display. This is the place where users can load their own python modules and casa tasks. For example `./casa/init.py` might contain:

```
import os
sys.path.insert(1,os.environ['HOME']+os.sep+"python")
import analysisUtils as aU
```

and `analysisUtils.py` might contain:

```
import numpy as np
from mpfit import mpfit
from pylab import *
from numpy.fft import fft
from scipy import polyfit
import taskinit as ti
from importasdm import importasdm
```

Many options can also be set in the file

```
~/casarc
```

E.g.

```
#
# Set these so that bug(), ask(), etc. know who you are
#
userinfo.name: Sheila User
userinfo.email: suser@nrao.edu
userinfo.org: NRAO

#NOTE: Fill this value in as appropriate - the units are MB
#It is important that you not set this value larger than your actual
#physical memory
#system.resources.memory: 2000
#help.popup.type: mb3long

#catalog
catalog.gui.auto: T
catalog.confirm: T
catalog.view.PostScript: ghostview
catalog.edit.ascii: xterm -e vi

#logger
#logger.file: ./aips++.log
#logger.height: 12
logger.default: screen

#progress meter GUI pop-ups - disable
progress.show: F

#toolmanager - disable
toolmanager.gui.auto: F

#Use current working directory for cache/scratch files
user.aipsdir: .
user.cache: .
user.directories.work: .
user.initfiles: almainit.g

#viewer
display.axislabels: on
display.colormaps.defaultcolormap: Hot Metal 1

#development

#ms.async: ddd ./ms %s
```

A.5 Updating the data repository

Each CASA release for linux comes with an up to date data repository (containing information such as observatory coordinates, calibrator models, leap second tables, etc.). However, the files that make up the data repository are updated regularly. Therefore, if you install (or have installed) a release that is a few weeks to a month old, it makes sense to update the data repository because it is very easy.

To do so, issue the following command from the CASA prompt:

```
CASA <2>: !update-data
```

Appendix B

Python and CASA

CASA uses Python, IPython and matplotlib within the package. IPython is an enhanced, interactive shell to Python which provides many features for efficient command line interaction, while matplotlib is a Python 2-D plotting library for publication quality figures in different hardcopy formats.

From www.python.org: "Python is an interpreted, interactive, object-oriented programming language". Python is used as the underlying command line interface/scripting language to CASA. Thus, CASA inherits the features and the annoyances of Python. For example, since Python is inherently 0-based in its indexing of arrays, vectors, etc, CASA is also 0-based; any Index inputs (e.g., start (for start channel), fieldIndex, antennaID, etc) will start with 0. Another example is that indenting of lines means something to Python, of which users will have to be aware.

Currently, CASA uses python 2.6 (2.5 for the Mac OS 10.5 version). Some key links to python are:

- <http://python.org> – Main Python page
- <https://docs.python.org/2/reference/> – Python Reference
- <https://docs.python.org/2/tutorial/> – Python Tutorial
- <http://ipython.scipy.org> – IPython page
- <http://matplotlib.sourceforge.net> – matplotlib page

Each of the features of these components behave in the standard way within CASA. In the following sections, we outline the key elements for analysis interactions; see the Python references and the IPython page for the full suite of functionality.

B.1 Python Packages

The following python packages are included in CASA: ipython, nose, pyfits, pytz, dbus, numpy, scientific python, twisted, zope.interface, foolscap, matplotlib, scipy. For their version numbers, please

check the CASA page <https://safe.nrao.edu/wiki/bin/view/Software/CasaDevelopmentDependencies>.

B.2 Automatic parentheses

Automatic parenthesis is enabled for calling functions with argument lists; this feature is intended to allow less typing for common situations. IPython will display the interpretation of the line, beneath the one typed, as indicated by the '----->'. Default behavior in CASA is to have automatic parenthesis enabled.

B.3 Indentation

Python pays attention to indentation of lines in scripts or when you enter them interactively. It uses indentation to determine the level of nesting in loops. Be careful when cutting and pasting, if you get the wrong indentation, then unpredictable things can happen (usually it just gives an error).

A blank line can be used to return the indentation to a previous level. For example, expanded parameters in tasks cause indentation in subsequent lines in the interface. For example, the following snippet of inputs from `clean` can be cut and pasted without error due to the blank line after the indented parameters:

```
mode           = 'channel'      # Type of selection
  nchan        =      -1        # Number of channels to select
  start        =         0      # Start channel
  step         =         1      # Increment between channels/velocity
  width        =         1      # Channel width

alg            = 'clark'        # Algorithm to use
```

If the blank line were not there, an error would result if you pasted this at the `casapy` prompt.

B.4 Lists and Ranges

Sometimes, you need to give a task a list of indices. For example, some tasks and tools expect a comma-separated Python list, e.g.

```
scanlist = [241, 242, 243, 244, 245, 246]
```

You can use the Python `range` function to generate a list of consecutive numbers, e.g.

```
scanlist = range(241,247)
```

giving the same list as above, e.g.


```
CASA <1>: scanlist=range(241,247)
CASA <2>: print scanlist
[241, 242, 243, 244, 245, 246]
```

Note that `range` starts from the first limit and goes to one below the second limit (Python is 0-based, and `range` is designed to work in loop functions). If only a single limit is given, the first limit is treated as 0, and the one given is used as the second, e.g.

```
CASA <3>: iflist=range(4)
CASA <4>: print iflist
[0, 1, 2, 3]
```

You can also combine multiple ranges by summing lists

```
CASA <5>: scanlist=range(241,247) + range(251,255)
CASA <6>: print scanlist
[241, 242, 243, 244, 245, 246, 251, 252, 253, 254]
```

B.5 Dictionaries

Python dictionaries are data structures that contain `key:value` pairs, sort of like a hash array. These are useful to store mini-databases of things. In CASA, the parameter values are kept in a dictionary behind the scenes.

To initialize a dictionary, say we call it `mydict`, for use:

```
CASA <7>: mydict = {}
```

To add members:

```
CASA <8>: mydict['source'] = '0137+331'
CASA <9>: mydict['flux'] = 5.4
```

To see its contents:

```
CASA <10>: mydict
Out[10]: {'flux': 5.4000000000000004, 'source': '0137+331'}
CASA <11>: print mydict
{'source': '0137+331', 'flux': 5.4000000000000004}
```

To access a specific entry:

```
CASA <12>: print mydict['flux']
5.4
```

B.5.1 Saving and Reading Dictionaries

To save a simple dictionary to a file:

```
CASA <13>: dictfile = open('mydictfile.py','w')
CASA <14>: print >>dictfile,"mydict = ",mydict
CASA <15>: dictfile.close()
CASA <16>: !cat mydictfile.py
IPython system call: cat mydictfile.py
mydict = {'source': '0137+331', 'flux': 5.4000000000000004}

CASA <17>: mydict = {}
CASA <18>: run mydictfile.py
CASA <19>: mydict
  Out[19]: {'flux': 5.4000000000000004, 'source': '0137+331'}
```

More complex dictionaries, like those produced by `imstat` that contain NumPy arrays, require a different approach to save. The `pickle` module lets you save general data structures from Python. For example:

```
CASA <20>: import pickle
CASA <21>: xstat
  Out[21]:
{'blc': array([0, 0, 0, 0]),
 'blcf': '15:24:08.404, +04.31.59.181, I, 1.41281e+09Hz',
 'flux': array([ 4.0795296]),
 'max': array([ 0.05235516]),
 'maxpos': array([134, 134,  0, 38]),
 'maxposf': '15:21:53.976, +05.05.29.998, I, 1.41374e+09Hz',
 'mean': array([ 1.60097857e-05]),
 'medabsdevmed': array([ 0.00127436]),
 'median': array([ -1.17422514e-05]),
 'min': array([-0.0104834]),
 'minpos': array([160,  1,  0, 30]),
 'minposf': '15:21:27.899, +04.32.14.923, I, 1.41354e+09Hz',
 'npts': array([ 3014656.]),
 'quartile': array([ 0.00254881]),
 'rms': array([ 0.00202226]),
 'sigma': array([ 0.0020222]),
 'sum': array([ 48.26399646]),
 'sumsq': array([ 12.32857318]),
 'trc': array([255, 255,  0, 45]),
 'trcf': '15:19:52.390, +05.35.44.246, I, 1.41391e+09Hz'}

CASA <22>: mydict
  Out[22]: {'flux': 5.4000000000000004, 'source': '0137+331'}

CASA <23>: pickfile = 'myxstat.pickle'
CASA <24>: f = open(pickfile,'w')
CASA <25>: p = pickle.Pickler(f)
```

```
CASA <26>: p.dump(xstat)
CASA <27>: p.dump(mydict)
CASA <28>: f.close()
```

The dictionaries are now saved in pickle file `myxstat.pickle` in the current directory.

To retrieve:

```
CASA <29>: xstat2 = {}
CASA <30>: mydict2 = {}
CASA <31>: f = open(pickfile)
CASA <32>: u = pickle.Unpickler(f)
CASA <33>: xstat2 = u.load()
CASA <34>: mydict2 = u.load()
CASA <35>: f.close()
CASA <36>: xstat2
  Out[36]:
{'blc': array([0, 0, 0, 0]),
 'blcf': '15:24:08.404, +04.31.59.181, I, 1.41281e+09Hz',
 'flux': array([ 4.0795296]),
 'max': array([ 0.05235516]),
 'maxpos': array([134, 134,  0, 38]),
 'maxposf': '15:21:53.976, +05.05.29.998, I, 1.41374e+09Hz',
 'mean': array([ 1.60097857e-05]),
 'medabsdevmed': array([ 0.00127436]),
 'median': array([ -1.17422514e-05]),
 'min': array([-0.0104834]),
 'minpos': array([160,  1,  0, 30]),
 'minposf': '15:21:27.899, +04.32.14.923, I, 1.41354e+09Hz',
 'npts': array([ 3014656.]),
 'quartile': array([ 0.00254881]),
 'rms': array([ 0.00202226]),
 'sigma': array([ 0.0020222]),
 'sum': array([ 48.26399646]),
 'sumsq': array([ 12.32857318]),
 'trc': array([255, 255,  0, 45]),
 'trcf': '15:19:52.390, +05.35.44.246, I, 1.41391e+09Hz'}
```

```
CASA <37>: mydict2
  Out[37]: {'flux': 5.4000000000000004, 'source': '0137+331'}
```

Thus, you can make scripts that save information and use it later, like for regressions.

Note that these examples use Python file-handling and IO, as well as importing modules such as `pickle`. See your friendly Python reference for more on this kind of stuff. It's fairly obvious how it works.

B.6 Control Flow: Conditionals, Loops, and Exceptions

There are a number of ways to control the flow of execution in Python, including conditionals (`if`), loops (`for` and `while`), and exceptions (`try`). We will discuss the first two below.

B.6.1 Conditionals

The standard `if` block handles conditional execution or branches in Python:

```
if <expression>:
    <statements>
elif <expression>:
    <statements>
elif <expression>:
    <statements>
...
else:
    <statements>
```

Insert a `pass` statement if you want no action to be taken for a particular clause. The `<expression>` should reduce down to `True` or `False`.

For example,

```
if ( importmode == 'vla' ):
    # Import the data from VLA Export to MS
    default('importvla')
    print "Use importvla to read VLA Export and make an MS"

    archivefiles = datafile
    vis = msfile
    bandname = exportband
    autocorr = False
    antnamescheme = 'new'
    project = exportproject

    importvla()
elif ( importmode == 'fits' ):
    # Import the data from VLA Export to MS
    default('importuvfits')
    print "Use importuvfits to read UVFITS and make an MS"

    fitsfile = datafile
    vis = msfile

    importuvfits()
else:
```

```
# Copy from msfile
print "Copying "+datafile+" to "+msfile
os.system('cp -r '+datafile+' '+msfile)
vis = msfile
```

chooses branches based on the value of the `importmode` Python variable (set previously in script).

B.6.2 Loops

The for loop

```
for iter in seq:
    <statements>
```

iterates over elements of a sequence `seq`, assigning each in turn to `iter`. The sequence is usually a list of values.

For example,

```
splitms = 'polcal_20080224.cband.all.split.ms'
srclist = ['0137+331', '2136+006', '2202+422', '2253+161', '0319+415', '0359+509']
spwlist = ['0', '1']
```

```
for src in srclist:
    for spwid in spwlist:
        imname = splitms + '.' + src + '.' + spwid + '.clean'
        clean(vis=splitms, field=src, spw=spwid, imagename=imname,
              stokes='IQUV', psfmode='hogbom', imagermode='csclean',
              imsize=[288,288], cell=[0.4,0.4], niter=1000,
              threshold=1.3, mask=[134,134,154,154])
```

```
# Done with spw
```

```
# Done with sources
```

As usual, blocks are closed by blank lines of the previous indentation level.

You can use the `range` (§ B.4) Python function to generate a numerical loop:

```
vis = 'polcal_20080224.cband.all.ms'
for i in range(0,6):
    fld = str(i)
    plotxy(vis, field=fld, xaxis='uvdist', yaxis='amp')
```

```
# Done with fields [0, 1, 2, 3, 4, 5]
```

There is also a `while` loop construct

```
while <expression>:
    <statements>
```

which executes the statement block while the `<expression>` is `True`. The `while` loop can also take an `else` block.

For example,

```
# Do an explicit set of clean iterations down to a limit
prevrms = 1.e10
while rms > 0.001 :
    clean(vis=splitms,field=src,spw=spwid,imagename=imname,
          stokes='IQUV',psfmode='hogbom',imagermode='csclean',
          imsize=[288,288],cell=[0.4,0.4],niter=200,
          threshold=1.3,mask=[134,134,154,154])

    offstat=imstat(imname+'.residual',box='224,224,284,284')
    rms=offstat['sigma'][0]
    if rms > prevrms:
        break                # the rms has increased, stop

    prevrms = rms

# Clean until the off-source rms residual, reaches 0.001 Jy
```

Note that you can exit a loop using the `break` statement, as we have here when the rms increases.

B.7 System shell access

For scripts, the `os.system` methods are the preferred way to access system shell commands (see § B.7.1).

In interactive mode, any input line beginning with a `'!` character is passed verbatim (minus the `'!`) to the underlying operating system. Several common commands (`ls`, `pwd`, `less`) may be executed with or without the `'!`. Note that the `cd` command must be executed without the `'!`, and the `cp` command must use `'!'` as there is a conflict with the `cp` tool in `casapy`.

For example:

```
CASA [1]: pwd
/export/home/corsair-vml/jmcmulli/data
CASA [2]: ls n*
ngc5921.ms ngc5921.py
CASA [3]: !cp -r ../test.py .
```

B.7.1 Using the `os.system` methods

To use this, you need the `os` package. This should be loaded by default by `casapy`, but if not you can use

```
import os
```

in your script.

For example, in our scripts we use this to clean up any existing output files

```
# The prefix to use for all output files
prefix='ngc5921.usecase'

# Clean up old files
os.system('rm -rf '+prefix+'*')
```

Note that the `os` package has many useful methods. You can see these by using tab-completion:

```
CASA <2>: os.<tab>
os.EX_CANTCREAT      os._Environ          os.fdatasync         os.remove
os.EX_CONFIG         os.__all__           os.fdopen            os.removedirs
os.EX_DATAERR        os.__builtins__     os.fork              os.rename
os.EX_IOERR          os.__class__        os.forkpty           os.renames
os.EX_NOHOST         os.__delattr__      os.fpathconf        os.rmdir
os.EX_NOINPUT        os.__dict__         os.fstat             os.sep
os.EX_NOPERM         os.__doc__          os.fstatvfs          os.setegid
os.EX_NOUSER         os.__file__         os.fsync             os.seteuid
os.EX_OK             os.__getattr__      os.ftruncate        os.setgid
os.EX_OSERR          os.__hash__         os.getcwd            os.setgroups
os.EX_OSFILE         os.__init__         os.getcwdu          os.setpgid
os.EX_PROTOCOL       os.__name__         os.getegid           os.setpgrp
os.EX_SOFTWARE       os.__new__          os.getenv            os.setregid
os.EX_TEMPFAIL       os.__reduce__       os.geteuid           os.setreuid
os.EX_UNAVAILABLE    os.__reduce_ex__   os.getgid            os.setsid
os.EX_USAGE          os.__repr__         os.getgroups         os.setuid
os.F_OK              os.__setattr__     os.getloadavg       os.spawnl
os.NGROUPS_MAX       os.__str__          os.getlogin         os.spawnle
os.O_APPEND          os._copy_reg        os.getpgid          os.spawnlp
os.O_CREAT           os._execvpe         os.getpgrp          os.spawnlpe
os.O_DIRECT          os._exists          os.getpid           os.spawnv
os.O_DIRECTORY       os._exit            os.getppid         os.spawnve
os.O_DSYNC           os._get_exports_list os.getsid           os.spawnvp
os.O_EXCL            os._make_stat_result os.getuid           os.spawnvpe
os.O_LARGEFILE       os._make_statvfs_result os.isatty          os.stat
os.O_NDELAY          os._pickle_stat_result os.kill             os.stat_float_times
os.O_NOCTTY          os._pickle_statvfs_result os.killpg          os.stat_result
os.O_NOFOLLOW        os._spawnvef        os.lchown           os.statvfs
os.O_NONBLOCK        os.abort            os.linesep          os.statvfs_result
os.O_RDONLY          os.access           os.link             os.strerror
```

os.O_RDONLY	os.altsep	os.listdir	os.symlink
os.O_RSYNC	os.chdir	os.lseek	os.sys
os.O_SYNC	os.chmod	os.lstat	os.sysconf
os.O_TRUNC	os.chown	os.major	os.sysconf_names
os.O_WRONLY	os.chroot	os.makedev	os.system
os.P_NOWAIT	os.close	os.makedirs	os.tcgetpgrp
os.P_NOWAITO	os.confstr	os.minor	os.tcsetpgrp
os.P_WAIT	os.confstr_names	os.mkdir	os.tempnam
os.R_OK	os.ctermid	os.mkfifo	os.times
os.SEEK_CUR	os.curdir	os.mknod	os.tmpfile
os.SEEK_END	os.defpath	os.name	os.tmpnam
os.SEEK_SET	os.devnull	os.nice	os.ttyname
os.TMP_MAX	os.dup	os.open	os.umask
os.UserDict	os.dup2	os.openpty	os.uname
os.WCONTINUED	os.environ	os.pardir	os.unlink
os.WCOREDUMP	os.errno	os.path	os.unsetenv
os.WEXITSTATUS	os.error	os.pathconf	os.urandom
os.WIFCONTINUED	os.execl	os.pathconf_names	os.utime
os.WIFEXITED	os.execlp	os.pathsep	os.wait
os.WIFSIGNALED	os.execlpe	os.pipe	os.wait3
os.WIFSTOPPED	os.execv	os.popen	os.wait4
os.WNOHANG	os.execve	os.popen2	os.waitpid
os.WSTOPSIG	os.execvp	os.popen3	os.walk
os.WTERMSIG	os.execvpe	os.popen4	os.write
os.WUNTRACED	os.extsep	os.putenv	
os.W_OK	os.fchdir	os.read	
os.X_OK		os.readlink	

B.7.2 Directory Navigation

In addition, filesystem navigation is aided through the use of bookmarks to simplify access to frequently-used directories:

```
CASA [4]: cd /home/ballista/jmcmulli/other_data
CASA [4]: pwd
/home/ballista/jmcmulli/other_data
CASA [5]: bookmark other_data
CASA [6]: cd /export/home/corsair-vml/jmcmulli/data
CASA [7]: pwd
/export/home/corsair-vml/jmcmulli/data
CASA [8]: cd -b other_data
(bookmark:data) -> /home/ballista/jmcmulli/other_data
```

For python scripts, there is a special command to change a directory.

```
os.system('cd ~/directory')
```

will NOT work but the following will:


```
os.chdir('~directory')
```

B.7.3 Shell Command and Capture

See also § B.9 for the use of the command history.

1. `sx shell_command`, `!!shell_command` - this captures the output to a list

```
CASA [1]: sx pwd # stores output of 'pwd' in a list
Out[1]: ['/home/basho3/jmcmulli/pretest']
```

```
CASA [2]: !!pwd # !! is a shortcut for 'sx'
Out[2]: ['/home/basho3/jmcmulli/pretest']
```

```
CASA [3]: sx ls v* # stores output of 'pwd' in a list
Out[3]:
['vla_calplot.jpg',
'vla_calplot.png',
'vla_msplot_cals.jpg',
'vla_msplot_cals.png',
'vla_plotcal_bpass.jpg',
'vla_plotcal_bpass.png',
'vla_plotcal_fcal.jpg',
'vla_plotcal_fcal.png',
'vla_plotvis.jpg',
'vla_plotvis.png']
```

```
CASA [4]: x=_ # remember '_' is a shortcut for the output from the last command
```

```
CASA [5]: x
Out[5]:
['vla_calplot.jpg',
'vla_calplot.png',
'vla_msplot_cals.jpg',
'vla_msplot_cals.png',
'vla_plotcal_bpass.jpg',
'vla_plotcal_bpass.png', 'vla_plotcal_fcal.jpg',
'vla_plotcal_fcal.png',
'vla_plotvis.jpg',
'vla_plotvis.png']
```

```
CASA [6]: y=Out[2] # or just refer to the enumerated output
```

```
CASA [7]: y
Out[7]: ['/home/basho3/jmcmulli/pretest']
```

2. `sc` - captures the output to a variable; options are `'-l'` and `'-v'`

```
CASA [1]: sc x=pwd # capture output from 'pwd' to the variable 'x'
```

```
CASA [2]: x
Out[2]: '/home/basho3/jmcmulli/pretest'

CASA [3]: sc -l x=pwd # capture the output from 'pwd' to the variable 'x' but
           # split newlines into a list (similar to sx command)

CASA [4]: x
Out[4]: ['/home/basho3/jmcmulli/pretest']

CASA [5]: sc -v x=pwd # capture output from 'pwd' to a variable 'x' and
           # show what you get (verbose mode)
x ==
'/home/basho3/jmcmulli/pretest'

CASA [6]: x
Out[6]: '/home/basho3/jmcmulli/pretest'
```

B.8 Logging

There are two components to logging within CASA. Logging of all command line inputs is done via IPython.

Upon startup, CASA will log all commands to a file called `ipython.log`. This file can be changed via the use of the `./casa/ipython/ipythonrc` file. This log file can be edited and re-executed as appropriate using the `execfile` feature (§ B.12).

Logging can be turned on and off using the `logon`, `logoff` commands.

The second component is the output from applications which is directed to the file `./casapy.log`. See § 1.5.2 for more on the `casalogger`.

B.9 History and Searching

Numbered input/output history is provided natively within IPython. Command history is also maintained on-line.

```
CASA [11]: x=1

CASA [12]: y=3*x

CASA [13]: z=x**2+y**2

CASA [14]: x
Out[14]: 1
```

```

CASA [15]: y
Out[15]: 3

CASA [16]: z
Out[16]: 10

CASA [17]: Out[14] # Note: The 'Out' vector contains command output
Out[17]: 1

CASA [18]: _15 # Note: The return value can be accessed by _number
Out[18]: 3

CASA [19]: ___ # Note: The last three return values can be accessed as:
Out[19]: 10 # -, __, ___

```

Command history can be accessed via the 'hist' command. The history is reset at the beginning of every CASA session, that is, typing 'hist' when you first start CASA will not provide any commands from the previous session. However, all of the commands are still available at the command line and can be accessed through the up or down arrow keys, and through searching.

```

CASA [22]: hist
1 : __IP.system("vi temp.py") # Note:shell commands are designated in this way
2 : ipmagic("run -i temp.py") # Note:magic commands are designated in this way
3 : ipmagic("hist ")
4 : more temp.py
5 : __IP.system("more temp.py")
6 : quickhelp() # Note: autoparenthesis are added in the history
7 : im.open('ngc5921.ms')
8 : im.summary()
9 : ipmagic("pdoc im.setdata")
10: im.close()
11: quickhelp()
12: ipmagic("logstate ")
13: x=1
14: y=3*x
15: z=x**2+y**2
16: x
17: y
18: z
19: Out[16]
20: _17
21: ___

```

The history can be saved as a script or used as a macro for further use:

```

CASA [24]: save script.py 13:16
File 'script.py' exists. Overwrite (y/[N])? y
The following commands were written to file 'script.py':
x=1

```

```

y=3*x
z=x**2+y**2
CASA [25]: !more script.py
x=1
y=3*x
z=x**2+y**2

```

Note that the history commands will be saved up to, but not including the last value (i.e., history commands 13-16 saves commands 13, 14, and 15).

There are two mechanisms for searching command history:

1. Previous/Next: use `Ctrl-p` (previous,up) and `Ctrl-n` (next,down) to search through only the history items that match what you have typed so far (min-match completion). If you use `Ctrl-p` or `Ctrl-n` at a blank prompt, they behave just like the normal arrow keys.
2. Search: `Ctrl-r` opens a search prompt. Begin typing and the system searches your history for lines that contain what you've typed so far, completing what it can. For example:

```

CASA [37]: <CTRL-r>

(reverse-i-search) ' ':

```

Typing anything after the colon will provide you with the last command matching the characters, for example, typing 'op' finds:

```

(reverse-i-search) 'op': im.open('ngc5921.ms')

```

Subsequent hitting of `Ctrl-r` will search for the next command matching the characters.

B.10 Macros

Macros can be made for easy re-execution of previous commands. For example to store the commands 13-15 to the macro 'example':

```

CASA [31]: macro example 13:16
Macro 'example' created. To execute, type its name (without quotes).
Macro contents:
x=1
y=3*x
z=x**2+y**2

CASA [32]: z
Out[32]: 6

CASA [33]: z=10

```

```
CASA [34]: example
Out[34]: Executing Macro...

CASA [35]: z
Out[35]: 6

CASA [36]:
```

B.11 On-line editing

You can edit files on-line in two ways:

1. Using the shell access via '!vi'
2. Using the ed function; this will edit the file but upon closing, it will try to execute the file; using the 'script.py' example above:

```
CASA [13]: ed script.py # this will bring up the file in your chosen editor
           # when you are finished editing the file,
           # it will automatically
           # execute it (as though you had done a
           # execfile 'script.py'
           Editing... done. Executing edited code...

CASA [14]: x
Out[14]: 1

CASA [15]: y
Out[15]: 3

CASA [16]: z
Out[16]: 6
```

B.12 Executing Python scripts

Python scripts are simple text files containing lists of commands as if typed at the keyboard. Note: the auto-parentheses feature of IPython cannot be used in scripts, that is, you should make sure all function calls have any opening and closing parentheses.

```
# file is script.py
# My script to plot the observed visibilities
plotxy('ngc5921.ms', 'uvdist') #yaxis defaults to amplitude
```

This can be done by using the execfile command to execute this script. execfile will execute the script as though you had typed the lines at the CASA prompt.

```
CASA [5]: execfile 'script.py'
-----> execfile('script.py')
```

If you don't want to launch CASA and execute your script from the *command line*, you can use the '-c' option:

```
unix$ casapy -c 'script.py'
```

B.13 How do I exit from CASA?

You can exit CASA by using the quit command. This will bring up the query

```
Do you really want to exit ([y]/n)?
```

to give you a chance in case you did not mean to exit. You can also quit using %exit or CTRL-D.

If you don't want to see the question "Do you really want to exit [y]/n?", then just type `Exit` or `exit` followed by return, and CASA will stop right then and there.

Appendix C

Appendix: Models, Conventions, and Reference Frames

This appendix lists the available parameters, conventions, reference frames, and information on flux standards used in CASA.

C.1 Flux Density Models for `setjy`

`setjy` adds a source model given the source name, frequency, a standard (really, a set of models), and possibly a time. At cm wavelengths the flux density (FD) calibrators are typically one of several bright extragalactic sources. These objects are comparatively faint and less well characterized at shorter wavelengths, so for (sub)mm astronomy it is common to use Solar System objects.

Reliably setting the FD scale with astronomical calibrators requires that they be bright, not too resolved, and have simple dependencies on frequency and time. These criteria are somewhat mutually exclusive, so the number of calibrator sources supported by `setjy` is fairly small, although it could certainly be added to. This appendix is for describing the models that `setjy` uses. Choosing a FD calibrator of course has to be done before the observation and the observatory may provide additional information.

C.1.1 Long wavelength calibration

Synchrotron sources can vary over a light crossing time, so ones used as FD calibrators must have most of their emission coming from an extended region. The additional requirement that they be nearly unresolved therefore forces them to be distant, meaning that candidates which also have high apparent fluxes are quite rare. The following standards mostly share the same set of objects, and monitor their FDs every few years to account for variations. No interpolation is done between epochs, though - you are encouraged to choose the standard which observed your FD calibrator closest to the time you observed it at. The measurements are interpolated in frequency, however, using second to fourth degree polynomials of the frequency's logarithm.

Table C.1: Recognized Flux Density Calibrators. Note that the VLA uses J2000 calibrator names. CASA accepts all strings that contain the names below. E.g. 'PKS 1934-638' will be recognized

3C Name	B1950 Name	J2000 Name	Alt. J2000 Name	Standards
3C48	0134+329	0137+331	J0137+3309	1,3,4,5,6, 7
3C123	0433+295	0437+296	J0437+2940	2
3C138	0518+165	0521+166	J0521+1638	1,3,4,5,6
3C147	0538+498	0542+498	J0542+4951	1,3,4,5,6, 7
3C196	0809+483	0813+482	J0813+4813	1,2,7
3C286	1328+307	1331+305	J1331+3030	1,2,3,4,5,6, 7
3C295	1409+524	1411+522	J1411+5212	1,2,3,4,5,6, 7
–	1934-638	–	J1939-6342	1,3,4,5,6
3C380	1828+487	1829+487	J1829+4845	7

Standards are: (1) Perley-Butler 2010, (2) Perley-Butler 2013, (3) Perley-Taylor 99, (4) Perley-Taylor 95, (5) Perley 90, (6) Baars, (7) Scaife-Heald 2012

C.1.1.1 Baars

The only standard to not have the year in the name. It is **1977**.

The models are second order polynomials in $\log(\nu)$, valid between 408 MHz and 15 GHz.

The paper is Baars, J. W. M., Genzel, R., Pauliny-Toth, I. I. K., & Witzel, A. 1977, A&A, 61, 99 with a commentary by Kellermann, K. I. 1999, A&A 500, 143.

C.1.1.2 Perley 90

This standard also includes 1934-638 from Reynolds (7/94) and 3C138 from Baars, J. W. M., Genzel, R., Pauliny-Toth, I. I. K., & Witzel, A. 1977, A&A, 61, 99.

Details can be found at <http://www.vla.nrao.edu/astro/calib/manual/baars.html>.

C.1.1.3 Perley-Taylor 95

Perley and Taylor (1995.2); plus Reynolds (1934-638; 7/94) Details can be found at <http://www.vla.nrao.edu/astro/calib/manual/baars.html>.

C.1.1.4 Perley-Taylor 99

Perley and Taylor (1999.2); plus Reynolds (1934-638; 7/94) Details can be found at <http://www.vla.nrao.edu/astro/calib/manual/baars.html>.

C.1.1.5 Perley-Butler 2010

A preliminary version of Perley-Butler 2013 (S C.1.1.6). This version has also coefficients for sources that showed some degree of variability, see Perley, R. A., & Butler, B. J. 2012, ApJS, submitted (<http://arxiv.org/abs/1211.1300>).

C.1.1.6 Perley-Butler 2013

Flux scale for the constant flux sources 3C123, 3C196, 3C286, and 3C295. The models are time-dependent.

Reference: Perley, R. A., & Butler, B. J. 2013, ApJS, 206, 16

C.1.1.7 Scaife-Heald 2012

Low frequency, 30-300MHz, calibrators 3C48, 3C147, 3C196, 3C286, 3C295, and 3C380.

Reference: Scaife, A. M., & Heald, G. H. 2012, MNRAS, 423, 30

C.1.2 Short wavelength calibration

The usual approach in this regime is to use models that are, to first order, thermal sources in the Solar System. Their apparent brightness of course varies in time with their distance from the Earth (and Sun), and orientation if they are not perfect spheres with zero obliquity. However, most of them have almost constant surface properties, so once those properties are measured their apparent brightness distributions can in principle be predicted for any time, given an ephemeris. Planets, in particular, however, have more complex spectra and effects such as atmospheric lines, magnetic fields, seasons, polar caps and surface features need to be taken into account when they are available and significant. In CASA the Solar System objects supported by `setjy` are available through the 'Butler-JPL-Horizons 2010', and 'Butler-JPL-Horizons 2012' standards. The models are described in ALMA Memo 594 available on https://science.nrao.edu/facilities/alma/aboutALMA/Technology/ALMA_Memo_Series/alma594/abs594.

C.2 Velocity Reference Frames

CASA supported velocity frames are listed in Table C.2.

C.2.1 Doppler Types

CASA supported Doppler types are listed in Table C.3.

Table C.2: Velocity frames in CASA

Name	Description
REST	Laboratory
LSRK	local standard of rest (kinematic)
LSRD	local standard of rest (dynamic)
BARY	barycentric
GEO	geocentric
TOPO	topocentric
GALACTO	galactocentric
LGROUP	Local Group
CMB	cosmic microwave background dipole
Undefined	undefined frame

Table C.3: Doppler types in CASA

Name	Description
RADIO	
Z	
RATIO	
BETA	
GAMMA	
OPTICAL	
TRUE	
RELATIVISTIC	

C.3 Time Reference Frames

CASA supported time reference frames are listed in Table C.4.

C.4 Coordinate Frames

CASA supported time coordinate frames are listed in Table C.5.

Note that TOPO refers to a time stamp at a given observation date. If more than one observation is concatenated this may lead to vastly erroneous values. Any conversion from TOPO to other frames such as BARY and LSRK should be performed for each individual observation, prior to concatenation or simultaneous imaging.

Table C.4: Time reference frames in CASA

Name	Description
LAST	
LMST	
GMST1	
GAST	
UT1	
UT2	
UTC	
TAI	
TDT	
TCG	
TDB	
TCB	
IAT	
GMST	
TT	
ET	
UT	

C.5 Physical Units

CASA also recognizes physical units. They are listed in Tables C.6, C.7, and C.8.

C.6 Physical Constants

The physical constants included in CASA can be found in Table C.9.

Table C.5: Coordinate frames in CASA

Name	Description
J2000	mean equator and equinox at J2000.0 (FK5)
JNAT	geocentric natural frame
JMEAN	mean equator and equinox at frame epoch
JTRUE	true equator and equinox at frame epoch
APP	apparent geocentric position
B1950	mean epoch and ecliptic at B1950.0.
B1950_VLA	mean epoch(1979.9) and ecliptic at B1950.0
BMEAN	mean equator and equinox at frame epoch
BTRUE	true equator and equinox at frame epoch
GALACTIC	Galactic coordinates
HADEC	topocentric HA and declination
AZEL	topocentric Azimuth and Elevation (N through E)
AZELSW	topocentric Azimuth and Elevation (S through W)
AZELNE	topocentric Azimuth and Elevation (N through E)
AZELGEO	geodetic Azimuth and Elevation (N through E)
AZELSWGEO	geodetic Azimuth and Elevation (S through W)
AZELNEGEO	geodetic Azimuth and Elevation (N through E)
ECLIPTC	ecliptic for J2000 equator and equinox
MECLIPTIC	ecliptic for mean equator of date
TECLIPTIC	ecliptic for true equator of date
SUPERGAL	supergalactic coordinates
ITRF	coordinates wrt ITRF Earth frame
TOPO	apparent topocentric position
ICRS	International Celestial reference system

Table C.6: Prefixes

Prefix	Name	Value
Y	(yotta)	10^{24}
Z	(zetta)	10^{21}
E	(exa)	10^{18}
P	(peta)	10^{15}
T	(tera)	10^{12}
G	(giga)	10^9
M	(mega)	10^6
k	(kilo)	10^3
h	(hecto)	10^2
da	(deka)	10
d	(deci)	10^{-1}
c	(centi)	10^{-2}
m	(milli)	10^{-3}
u	(micro)	10^{-6}
n	(nano)	10^{-9}
p	(pico)	10^{-12}
f	(femto)	10^{-15}
a	(atto)	10^{-18}
z	(zepto)	10^{-21}
y	(yocto)	10^{-24}

Table C.7: SI Units

Unit	Name	Value
\$	(currency)	1 -
%	(percent)	0.01
%%	(permille)	0.001
A	(ampere)	1 A
AE	(astronomical unit)	149597870659 m
AU	(astronomical unit)	149597870659 m
Bq	(becquerel)	1 s ⁻¹
C	(coulomb)	1 s A
F	(farad)	1 m ⁻² kg ⁻¹ s ⁴ A ²
Gy	(gray)	1 m ² s ⁻²
H	(henry)	1 m ² kg s ⁻² A ⁻²
Hz	(hertz)	1 s ⁻¹
J	(joule)	1 m ² kg s ⁻²
Jy	(jansky)	10 ⁻²⁶ kg s ⁻²
K	(kelvin)	1 K
L	(litre)	0.001 m ³
M0	(solar mass)	1.98891944407 × 10 ³⁰ kg
N	(newton)	1 m kg s ⁻²
Ohm	(ohm)	1 m ² kg s ⁻³ A ⁻²
Pa	(pascal)	1 m ⁻¹ kg s ⁻²
S	(siemens)	1 m ⁻² kg ⁻¹ s ³ A ²
S0	(solar mass)	1.98891944407 × 10 ³⁰ kg
Sv	(sievert)	1 m ² s ⁻²
T	(tesla)	1 kg s ⁻² A ⁻¹
UA	(astronomical unit)	149597870659 m
V	(volt)	1 m ² kg s ⁻³ A ⁻¹
W	(watt)	1 m ² kg s ⁻³
Wb	(weber)	1 m ² kg s ⁻² A ⁻¹
-	(undimensioned)	1 -

Table C.7: SI Units – continued

Unit	Name	Value
a	(year)	31557600 s
arcmin	(arcmin)	0.000290888208666 rad
arcsec	(arcsec)	$4.8481368111 \times 10^{-6}$ rad
as	(arcsec)	$4.8481368111 \times 10^{-6}$ rad
cd	(candela)	1 cd
cy	(century)	3155760000 s
d	(day)	86400 s
deg	(degree)	0.0174532925199 rad
g	(gram)	0.001 kg
h	(hour)	3600 s
l	(litre)	0.001 m ³
lm	(lumen)	1 cd sr
lx	(lux)	1 m ⁻² cd sr
m	(metre)	1 m
min	(minute)	60 s
mol	(mole)	1 mol
pc	(parsec)	$3.08567758065 \times 10^{16}$ m
rad	(radian)	1 rad
s	(second)	1 s
sr	(steradian)	1 sr
t	(tonne)	1000 kg

Table C.8: Custom Units

Unit	Name	Value
"	(arcsec)	$4.8481368111 \times 10^{-6}$ rad
"_2	(square arcsec)	$2.35044305391 \times 10^{-11}$ sr
'	(arcmin)	0.000290888208666 rad
"	(arcsec)	$4.8481368111 \times 10^{-6}$ rad
"_2	(square arcsec)	$2.35044305391 \times 10^{-11}$ sr
'_2	(square arcmin)	$8.46159499408 \times 10^{-8}$ sr
:	(hour)	3600 s
::	(minute)	60 s
:::	(second)	1 s
Ah	(ampere hour)	3600 s A
Angstrom	(angstrom)	1e-10 m
Btu	(British thermal unit (Int))	$1055.056 \text{ m}^2 \text{ kg s}^{-2}$
CM	(metric carat)	0.0002 kg
Cal	(large calorie (Int))	$4186.8 \text{ m}^2 \text{ kg s}^{-2}$
FU	(flux unit)	$10^{-26} \text{ kg s}^{-2}$
G	(gauss)	$0.0001 \text{ kg s}^{-2} \text{ A}^{-1}$
Gal	(gal)	0.01 m s^{-2}
Gb	(gilbert)	0.795774715459 A
Mx	(maxwell)	$10^{-8} \text{ m}^2 \text{ kg s}^{-2} \text{ A}^{-1}$
Oe	(oersted)	$79.5774715459 \text{ m}^{-1} \text{ A}$
R	(mile)	$0.000258 \text{ kg}^{-1} \text{ s A}$
St	(stokes)	$0.0001 \text{ m}^2 \text{ s}^{-1}$
Torr	(torr)	$133.322368421 \text{ m}^{-1} \text{ kg s}^{-2}$
USfl_oz	(fluid ounce (US))	$2.95735295625 \times 10^{-5} \text{ m}^3$
USgal	(gallon (US))	$0.003785411784 \text{ m}^3$

Table C.8: Custom Units – continued

Unit	Name	Value
WU	(WSRT flux unit)	$5 \times 10^{-29} \text{ kg s}^{-2}$
abA	(abampere)	10 A
abC	(abcoulomb)	10 s A
abF	(abfarad)	$10^9 \text{ m}^{-2} \text{ kg}^{-1} \text{ s}^4 \text{ A}^2$
abH	(abhenry)	$10^{-9} \text{ m}^2 \text{ kg s}^{-2} \text{ A}^{-2}$
abOhm	(abohm)	$10^{-9} \text{ m}^2 \text{ kg s}^{-3} \text{ A}^{-2}$
abV	(abvolt)	$10^{-8} \text{ m}^2 \text{ kg s}^{-3} \text{ A}^{-1}$
ac	(acre)	4046.8564224 m ²
arcmin_2	(square arcmin)	$8.46\text{-}2159499408 \times 10^{-8} \text{ sr}$
arcsec_2	(square arcsec)	$2.35044305391 \times 10^{-11} \text{ sr}$
ata	(technical atmosphere)	$98066.5 \text{ m}^{-1} \cdot \text{kg} \cdot \text{s}^{-2}$
atm	(standard atmosphere)	$101325 \text{ m}^{-1} \cdot \text{kg} \cdot \text{s}^{-2}$
bar	(bar)	$100000 \text{ m}^{-1} \cdot \text{kg} \cdot \text{s}^{-2}$
beam	(undefined beam area)	1 -
cal	(calorie (Int))	$4.1868 \text{ m}^2 \text{ kg s}^{-2}$
count	(count)	1 -
cwt	(hundredweight)	50.80234544 kg
deg_2	(square degree)	0.000304617419787 sr
dyn	(dyne)	$10^{-5} \text{ m kg s}^{-2}$
eV	(electron volt)	$1.60217733 \times 10^{-19} \text{ m}^2 \text{ kg s}^{-2}$
erg	(erg)	$10^{-7} \text{ m}^2 \text{ kg s}^{-2}$
fl_oz	(fluid ounce (Imp))	$2.84130488996 \times 10^{-5} \text{ m}^3$
ft	(foot)	0.3048 m
fu	(flux unit)	$10^{-26} \text{ kg s}^{-2}$
fur	(furlong)	201.168 m
gal	(gallon (Imp))	$0.00454608782394 \text{ m}^3$

Table C.8: Custom Units – continued

Unit	Name	Value
ha	(hectare)	10000 m ²
hp	(horsepower)	745.7 m ² kg s ⁻³
in	(inch)	0.0254 m
kn	(knot (Imp))	0.514773333333 m s ⁻¹
lambda	(lambda)	1 -
lb	(pound (avoirdupois))	0.45359237 kg
ly	(light year)	9.46073047 × 10 ¹⁵ m
mHg	(metre of mercury)	133322.387415 m ⁻¹ kg s ⁻²
mile	(mile)	1609.344 m
n_mile	(nautical mile (Imp))	1853.184 m
oz	(ounce (avoirdupois))	0.028349523125 kg
pixel	(pixel)	1 -
sb	(stilb)	10000 m ⁻² cd
sq_arcmin	(square arcmin)	8.46159499408 × 10 ⁻⁸ sr
sq_arcsec	(square arcsec)	2.35044305391 × 10 ⁻¹¹ sr
sq_deg	(square degree)	0.000304617419787 sr
statA	(statampere)	3.33564095198 × 10 ⁻¹⁰ A
statC	(statcoulomb)	3.33564095198 × 10 ⁻¹⁰ s A
statF	(statfarad)	1.11188031733 × 10 ⁻¹² m ⁻² kg ⁻¹ s ⁴ A ²
statH	(stathenry)	899377374000 m ² kg s ⁻² A ⁻²
statOhm	(statohm)	899377374000 m ² kg s ⁻³ A ⁻²
statV	(statvolt)	299.792458 m ² kg s ⁻³ A ⁻¹
u	(atomic mass unit)	1.661 × 10 ⁻²⁷ kg
yd	(yard)	0.9144 m
yr	(year)	31557600 s

Table C.9: Physical Constants

Constant	Name	Value
pi	3.14..	3.14159
ee	2.71..	2.71828
c	light vel.	$2.99792 \times 10^8 \text{ m s}^{-1}$
G	grav. const	$6.67259 \times 10^{-11} \text{ N m}^2 \text{ kg}^{-2}$
h	Planck const	$6.62608 \times 10^{-34} \text{ J s}$
HI	HI line	1420.41 MHz
R	gas const	$8.31451 \text{ J K}^{-1} \text{ mol}^{-1}$
NA	Avogadro #	$6.02214 \times 10^{23} \text{ mol}^{-1}$
e	electron charge	$1.60218 \times 10^{-19} \text{ C}$
mp	proton mass	$1.67262 \times 10^{-27} \text{ kg}$
mp_me	mp/me	1836.15
mu0	permeability vac.	$1.25664 \times 10^{-6} \text{ H m}^{-1}$
eps0	permittivity vac.	$1.60218 \times 10^{-19} \text{ C}$
k	Boltzmann const	$1.38066 \times 10^{-23} \text{ J K}^{-1}$
F	Faraday const	$96485.3 \text{ C mol}^{-1}$
me	electron mass	$9.10939 \times 10^{-31} \text{ kg}$
re	electron radius	$2.8179 \times 10^{-15} \text{ m}$
a0	Bohrs radius	$5.2918 \times 10^{-11} \text{ m}$
R0	solar radius	$6.9599 \times 10^8 \text{ m}$
k2	IAU grav. const ²	$0.000295912 \text{ AU}^3 \text{ d}^{-2} \text{ S}0^{-1}$

Appendix D

Appendix: CASA Region File Format

The CASA region file format provides a flexible, easily edited set of region definitions which are accepted across CASA tasks. Region files may be written by hand or using the *CASA viewer*.

Alert: Whereas the region format is supported by all the data processing tasks, the *viewer* implementation is still limited to rectangles, ellipses, and some markers.

For a file to be recognized as a valid CASA region text file, the first line must contain the string:

```
#CRTF
```

"CRTF" stands for "CASA Region Text Format". One may also include an optional version number at the end of the string, so it reads `#CRTFv0`; this indicates the version of the format definition.

Region files have two different kinds of definitions, "*regions*" and "*annotations*", each of which is one line long. To indicate an annotation, a line must begin with "`ann`". Lines that begin with the comment character (`#`) are not considered for processing or display.

The second line of a file may define global parameters that are to be used for all regions and annotations in that file, in which case the line starts with the word "`global`". The parameters set here may also be overridden by keywords in a specific line, in which case the keywords pertain only to that one line.

- **Regions:** all regions are considered by tasks. They will be displayed by visualization tasks as well as used to create masks, etc., as appropriate.
- **Annotations:** these are used by display tasks, and are for visual reference only.

D.1 Region definitions

All regions lines will follow this general arrangement:

{shape} {additional parameter=value pairs}

The possible parameter/value pairs are described in more detail below. Note that most parameters beyond the shape and its coordinates can be defined globally.

Possible units for coordinates are:

- *sexagesimal*, e.g. 18h12m24s for right ascension or -03.47.27.1 for declination
- *decimal degrees*, e.g. 140.0342deg for both RA and Dec
- *radians*, e.g. 2.37666rad for both RA and Dec
- *pixels*, e.g. 204pix

Possible units of length are:

- *degrees*, e.g. 23deg
- *arcminutes*, e.g. 23arcmin
- *arcseconds*, e.g. 23arcsec
- *radians*, e.g. 0.00035rad
- *pixels*, e.g. 23pix

Units must always be included when defining a region. If any fractional pixels are covered, the entire pixel will be selected. For purely single-pixel work there are alternate methods which may be preferable to using regions, eg. `ia.topixel`, `ia.toworld`, `ia.pixelvalue`.

D.2 Allowed shapes

- **Rectangular box**; the two coordinates are two opposite corners:

```
box[[x1, y1], [x2, y2]]
```

- **Center box**; `[x, y]` define the center point of the box and `[x_width, y_width]` the width of the sides:

```
centerbox[[x, y], [x_width, y_width]]
```

- **Rotated box**; `[x, y]` define the center point of the box; `[x_width, y_width]` the width of the sides; `rotang` the rotation angle:

```
rotbox[[x, y], [x_width, y_width], rotang]
```

- **Polygon**; there could be many $[x, y]$ corners; note that the last point will connect with the first point to close the polygon:

```
poly[[x1, y1], [x2, y2], [x3, y3], ...]
```

- *Circle*; center of the circle $[x,y]$, r is the radius:

```
circle[[x, y], r]
```

- *Annulus*; center of the circle is $[x, y]$, $[r1, r2]$ are inner and outer radii:

```
annulus[[x, y], [r1, r2]]
```

- *Ellipse*; center of the ellipse is $[x, y]$; semi-major and semi-minor axes are $[bmaj, bmin]$; position angle of the major axis is pa :

```
ellipse[[x, y], [bmaj, bmin], pa]
```

D.3 Annotation definitions

In addition to the definitions for regions [above], the following are always treated as annotations:

- *Line*; coordinates define the end points of the line:

```
line[[x1, y1], [x2, y2]]
```

- *Vector*; coordinates define end points; second coordinate pair is location of tip of arrow:

```
vector[[x1, y1], [x2, y2]]
```

- *Text*; coordinates define leftmost point of text string:

```
text[[x, y], 'my text']
```

- *Symbol*; coordinates define location of symbol (see Sec. D.7.1 for a list of allowed symbols):

```
symbol[[x, y], {symbol}]
```

D.4 Global definitions

Definitions to be used throughout the region file are placed on a line beginning with **'global'**, usually at the top of the file. These definitions may also be used on any individual region or annotation line; in this case, the value defined on that line will override the predefined global (but only for that line). If a *'global'* line occurs later in the file, subsequent lines will obey those definitions.

- *Coordinate reference frame:*

- Possible values: J2000, JMEAN, JTRUE, APP, B1950, B1950_VLA, BMEAN, BTRUE, GALACTIC, HADEC, AZEL, AZELSW, AZELNE, AZELGEO, AZELSWGEO, AZELNEGEO, JNAT, ECLIPTIC, MECLIPTIC, TECLIPTIC, SUPERGAL, ITRF, TOPO, ICRS

- Default: image value

```
coord = J2000
```

- *Frequency/velocity axis:*

- Possible values: REST, LSRK, LSRD, BARY, GEO, TOPO, GALACTO, LGROUP, CMB

- Default: image value

```
frame=TOPO
```

- *Frequency/velocity range:*

- Possible units: GHz, MHz, kHz, km/s, Hz, channel, chan (=channel)

- Default: image range

```
range=[min, max]
```

- *Correlation axis:*

- Possible values: I, Q, U, V, RR, RL, LR, LL, XX, XY, YX, YY, RX, RY, LX, LY, XR, XL, YR, YL, PP, PQ, QP, QQ, RCircular, LCircular, Linear, Ptotal, Plinear, PFtotal, PFlinear, Pangle

- Default: all planes present in image

```
corr=[X, Y]
```

- *Velocity calculation:*

- Possible values: RADIO, OPTICAL, Z, BETA, GAMMA

- Default: image value

```
veltype=RADIO
```

- *Rest frequency:*

- Default: image value

```
restfreq=1.42GHz
```

- *Line characteristics:*

- Possible values: any line style recognized by matplotlib: '-'=solid, '--'=dashed, ':'=dotted
- Default: linewidth=1, linestyle='-'

```
linewidth=1
linestyle='-'
```

- *Symbol characteristics:*

- Symbol size and thickness:

```
symsize = 1
symthick = 1
```

- *Region, symbol, and text color:*

- Possible values: any color recognized by matplotlib, including hex values
- Default: color=green

```
color=red
```

- *Text font characteristics:*

- Possible values: see Sect. D.7.2.
- 'usetex' is a boolean parameter that determines whether or not the text line should be interpreted as LaTeX, and would require working LaTeX, dvipng, and Ghostscript installations (equivalent to the text.usetex parameter in matplotlib).

```
font=Helvetica
fontsize=10pt
fontstyle=bold
usetex=True/False
```


- *Label position:*
 - Possible values: 'left', 'right', 'top', 'bottom'
 - Default: 'top'
 - `labelpos='right'`
- *Label color:*
 - Default: color of associated region.
 - Allowed values: same as values for region colors.
 - `labelcolor='green'`
- *Label offset:*
 - Default: [0,0].
 - Allowed values: any positive or negative number, in units of pixels.
 - `labeloff=[1, 1]`

D.5 Allowed additional parameters

These must be defined per region line:

- *Labels:* text label for a region; should be placed so text does not overlap with region boundary
 - `label='string'`
- *"OR/NOT" operators:* A "+" at the beginning of a line will flag it with a boolean "OR" (default), and a "-" will flag it with a boolean "NOT". Overlapping regions will be treated according to their sequence in the file; i.e., (((entireImage OR line1) OR line2) NOT line3) OR line4). This allows some flexibility in building "non-standard" regions. Note that a task (e.g., clean) will still consider all lines: if one wishes to remove a region from consideration, it should be commented out ("#").
- Default: OR (+)

D.6 Examples

A file with both global definitions and per-line definitions:

```
#CRTFv0
global coord=B1950_VLA, frame=BARY, corr=[I, Q], color=blue
```

```

# A simple circle region:
circle[[18h12m24s, -23d11m00s], 2.3arcsec]

# A box region, this one only for annotation:
ann box[[140.0342deg, -12.34243deg], [140.0360deg, -12.34320deg]]

# A rotated box region, for a particular range of velocities:
rotbox[[12h01m34.1s, 12d23m33s], [3arcmin, 1arcmin], 12deg], range=[-1240km/s, 1240km/s]

# An annular region, overriding some of the global defaults:
annulus[[17h51m03.2s, -45d17m50s], [0.10deg, 4.12deg]], corr=[I,Q,U,V], color=red,
label='My label here'

# Cuts an ellipse out of the previous regions, but only for Q and a particular frequency
range:
-ellipse[[17:51:03.2, -45.17.50], [0.25deg, 1.34deg], 45rad], range=[1.420GHz, 1.421GHz],
corr=[Q], color=green, label='Removed this'

# A diamond marker, in J2000 coordinates:
symbol[[32.1423deg, 12.1412deg], D], linewidth=2, coord=J2000, symsize=2

```

D.7 Fonts and Symbols

D.7.1 Allowed symbols

```

'.' point marker
',' pixel marker
'o' circle marker
'v' triangle_down marker
'^' triangle_up marker
'<' triangle_left marker
'>' triangle_right marker
'1' tri_down marker
'2' tri_up marker
'3' tri_left marker
'4' tri_right marker
's' square marker
'p' pentagon marker
'*' star marker
'h' hexagon1 marker
'H' hexagon2 marker
'+' plus marker
'x' x marker
'D' diamond marker
'd' thin_diamond marker
'|' vline marker

```

'_' hline marker

D.7.2 Allowed fonts

D.7.2.1 Allowed fonts for Linux

"Century Schoolbook L", "Console", "Courier", "Courier 10 Pitch", "Cursor", "David CLM", "DejaVu LGC Sans", "DejaVu LGC Sans Condensed", "DejaVu LGC Sans Light", "DejaVu LGC Sans Mono", "DejaVu LGC Serif", "DejaVu LGC Serif Condensed", "Dingbats", "Drugulin CLM", "East Syriac Adiabene", "Ellinia CLM", "Estrangelo Antioch", "Estrangelo Edessa", "Estrangelo Nisibin", "Estrangelo Nisibin Outline", "Estrangelo Talada", "Fangsong ti", "Fixed [Sony]", "Fixed [Eten]", "Fixed [Misc]", "Fixed [MNKANAME]", "Frank Ruehl CLM", "fxd", "Goha-Tibeb Zemen", "goth_p", "Gothic [Shinonome]", "Gothic [mplus]", "hlv", "hlvw", "KacstArt", "KacstBook", "KacstDecorative", "KacstDigital", "KacstFarsi", "KacstLetter", "KacstPoster", "KacstQura", "KacstQuraFixed", "KacstQuran", "KacstTitle", "KacstTitleL", "Liberation Mono", "Liberation Sans", "Liberation Serif", "LKLUG", "Lohit Bengali", "Lohit Gujarati", "Lohit Hindi", "Lohit Kannada", "Lohit Malayalam", "Lohit Oriya", "Lohit Punjabi", "Lohit Tamil", "Lohit Telugu", "LucidaTypewriter", "Luxi Mono", "Luxi Sans", "Luxi Serif", "Marumoji", "Miriam CLM", "Miriam Mono CLM", "MiscFixed", "Monospace", "Nachlieli CLM", "Nimbus Mono L", "Nimbus Roman No9 L", "Nimbus Sans L", "Nimbus Sans L Condensed", "PakTypeNaqsh", "PakTypeTehreer", "qub", "Sans Serif", "Sazanami Gothic", "Sazanami Mincho", "Serif", "Serto Batman", "Serto Jerusalem", "Serto Jerusalem Outline", "Serto Mardin", "Standard Symbols L", "sys", "URW Bookman L", "URW Chancery L", "URW Gothic L", "URW Palladio L", "Utopia", "Yehuda CLM",

D.7.2.2 Allowed fonts for MacOS X

"Abadi MT Condensed Light", "Adobe Caslon Pro", "Adobe Garamond Pro", "Al Bayan", "American Typewriter", "Andale Mono", "Apple Braille", "Apple Chancery", "Apple LiGothic", "Apple LiSung", "Apple Symbols", "AppleGothic", "AppleMyungjo", "Arial", "Arial Black", "Arial Hebrew", "Arial Narrow", "Arial Rounded MT Bold", "Arial Unicode MS", "Arno Pro", "Ayuthaya", "Baghdad", "Baskerville", "Baskerville Old Face", "Batang", "Bauhaus 93", "Bell Gothic Std", "Bell MT", "Bernard MT Condensed", "BiauKai", "Bickham Script Pro", "Big Caslon", "Birch Std", "Blackoak Std", "Book Antiqua", "Bookman Old Style", "Bookshelf Symbol 7", "Braggadocio", "Britannic Bold", "Brush Script MT", "Brush Script Std", "Calibri", "Calisto MT", "Cambria", "Candara", "Century", "Century Gothic", "Century Schoolbook", "Chalkboard", "Chalkduster", "Chaparral Pro", "Charcoal CY", "Charlemagne Std", "Cochin", "Colonna MT", "Comic Sans MS", "Consolas", "Constantia", "Cooper Black", "Cooper Std", "Copperplate", "Copperplate Gothic Bold", "Copperplate Gothic Light", "Corbel", "Corsiva Hebrew", "Courier", "Courier New", "Curlz MT", "DecoType Naskh", "Desdemona", "Devanagari MT", "Didot", "Eccentric Std", "Edwardian Script ITC", "Engravers MT", "Euphemia UCAS", "Eurostile", "Footlight MT Light", "Franklin Gothic Book", "Franklin Gothic Medium", "Futura", "Garamond", "Garamond Premier Pro", "GB18030 Bitmap", "Geeza Pro", "Geneva", "Geneva CY", "Georgia", "Giddyup

Std", "Gill Sans", "Gill Sans MT", "Gill Sans Ultra Bold", "Gloucester MT Extra Condensed", "Goudy Old Style", "Gujarati MT", "Gulim", "GungSeo", "Gurmukhi MT", "Haettenschweiler", "Harrington", "HeadLineA", "Hei", "Heiti SC", "Heiti TC", "Helvetica", "Helvetica CY", "Helvetica Neue", "Herculanum", "Hiragino Kaku Gothic Pro", "Hiragino Kaku Gothic ProN", "Hiragino Kaku Gothic Std", "Hiragino Kaku Gothic StdN", "Hiragino Maru Gothic Pro", "Hiragino Maru Gothic ProN", "Hiragino Mincho Pro", "Hiragino Mincho ProN", "Hiragino Sans GB", "Hobo Std", "Hoefler Text", "Impact", "Imprint MT Shadow", "InaiMathi", "Kai", "Kailasa", "Kino MT", "Kokonor", "Kozuka Gothic Pro", "Kozuka Mincho Pro", "Krungthep", "KufiStandardGK", "Letter Gothic Std", "LiHei Pro", "LiSong Pro", "Lithos Pro", "Lucida Blackletter", "Lucida Bright", "Lucida Calligraphy", "Lucida Console", "Lucida Fax", "Lucida Grande", "Lucida Handwriting", "Lucida Sans", "Lucida Sans Typewriter", "Lucida Sans Unicode", "Marker Felt", "Marlett", "Matura MT Script Capitals", "Meiryo", "Menlo", "Mesquite Std", "Microsoft Sans Serif", "Minion Pro", "Mistral", "Modern No. 20", "Monaco", "Monotype Corsiva", "Monotype Sorts", "MS Gothic", "MS Mincho", "MS PGothic", "MS PMincho", "MS Reference Sans Serif", "MS Reference Specialty", "Mshtakan", "MT Extra", "Myriad Pro", "Nadeem", "New Peninim MT", "News Gothic MT", "Nueva Std", "OCR A Std", "Onyx", "Optima", "Orator Std", "Osaka", "Papyrus", "PCMyungjo", "Perpetua", "Perpetua Titling MT", "PilGi", "Plantagenet Cherokee", "Playbill", "PMingLiU", "Poplar Std", "Prestige Elite Std", "Raana", "Rockwell", "Rockwell Extra Bold", "Rosewood Std", "Sathu", "Silom", "SimSun", "Skia", "Stencil", "Stencil Std", "STFangsong", "STHeiti", "STKaiti", "STSong", "Symbol", "Tahoma", "Tekton Pro", "Thonburi", "Times", "Times New Roman", "Trajan Pro", "Trebuchet MS", "Tw Cen MT", "Verdana", "Webdings", "Wide Latin", "Wingdings", "Wingdings 2", "Wingdings 3", "Zapf Dingbats", "Zapfino"

Appendix E

The Measurement Equation and Calibration

The visibilities measured by an interferometer must be calibrated before formation of an image. This is because the wavefronts received and processed by the observational hardware have been corrupted by a variety of effects. These include (but are not exclusive to): the effects of transmission through the atmosphere, the imperfect details amplified electronic (digital) signal and transmission through the signal processing system, and the effects of formation of the cross-power spectra by a correlator. Calibration is the process of reversing these effects to arrive at corrected visibilities which resemble as closely as possible the visibilities that would have been measured in vacuum by a perfect system. The subject of this chapter is the determination of these effects by using the visibility data itself.

E.1 The HBS Measurement Equation

The relationship between the observed and ideal (desired) visibilities on the baseline between antennas i and j may be expressed by the Hamaker-Bregman-Sault *Measurement Equation*¹:

$$\vec{V}_{ij} = J_{ij} \vec{V}_{ij}^{\text{IDEAL}}$$

where \vec{V}_{ij} represents the observed visibility, $\vec{V}_{ij}^{\text{IDEAL}}$ represents the corresponding ideal visibilities, and J_{ij} represents the accumulation of all corruptions affecting baseline ij . The visibilities are indicated as vectors spanning the four correlation combinations which can be formed from dual-polarization signals. These four correlations are related directly to the Stokes parameters which fully describe the radiation. The J_{ij} term is therefore a 4×4 matrix.

Most of the effects contained in J_{ij} (indeed, the most important of them) are antenna-based, i.e., they arise from measurable physical properties of (or above) individual antenna elements in a synthesis array. Thus, adequate calibration of an array of N_{ant} antennas forming $N_{ant}(N_{ant} - 1)/2$ baseline visibilities is usually achieved through the determination of only N_{ant} factors, such that

¹Hamaker, J.P., Bregman, J.D. & Sault, R.J. (1996), *Astronomy and Astrophysics Supplement*, v.117, p.137-147

$J_{ij} = J_i \otimes J_j^*$. For the rest of this chapter, we will usually assume that J_{ij} is factorable in this way, unless otherwise noted.

As implied above, J_{ij} may also be factored into the sequence of specific corrupting effects, each having their own particular (relative) importance and physical origin, which determines their unique algebra. Including the most commonly considered effects, the Measurement Equation can be written:

$$\vec{V}_{ij} = M_{ij} B_{ij} G_{ij} D_{ij} E_{ij} P_{ij} T_{ij} \vec{V}_{ij}^{\text{IDEAL}}$$

where:

- T_{ij} = Polarization-independent multiplicative effects introduced by the troposphere, such as opacity and path-length variation.
- P_{ij} = Parallactic angle, which describes the orientation of the polarization coordinates on the plane of the sky. This term varies according to the type of the antenna mount.
- E_{ij} = Effects introduced by properties of the optical components of the telescopes, such as the collecting area's dependence on elevation.
- D_{ij} = Instrumental polarization response. "D-terms" describe the polarization leakage between feeds (e.g. how much the R-polarized feed picked up L-polarized emission, and vice versa).
- G_{ij} = Electronic gain response due to components in the signal path between the feed and the correlator. This complex gain term G_{ij} includes the scale factor for absolute flux density calibration, and may include phase and amplitude corrections due to changes in the atmosphere (in lieu of T_{ij}). These gains are polarization-dependent.
- B_{ij} = Bandpass (frequency-dependent) response, such as that introduced by spectral filters in the electronic transmission system
- M_{ij} = Baseline-based correlator (non-closing) errors. By definition, these are not factorable into antenna-based parts.

Note that the terms are listed in the order in which they affect the incoming wavefront (G and B represent an arbitrary sequence of such terms depending upon the details of the particular electronic system). Note that M differs from all of the rest in that it is not antenna-based, and thus not factorable into terms for each antenna.

As written above, the measurement equation is very general; not all observations will require treatment of all effects, depending upon the desired dynamic range. E.g., bandpass need only be considered for continuum observations if observed in a channelized mode and very high dynamic range is desired. Similarly, instrumental polarization calibration can usually be omitted when observing (only) total intensity using circular feeds. Ultimately, however, each of these effects occurs at some level, and a complete treatment will yield the most accurate calibration. Modern high-sensitivity instruments such as ALMA and JVLA will likely require a more general calibration

treatment for similar observations with older arrays in order to reach the advertised dynamic ranges on strong sources.

In practice, it is usually far too difficult to adequately measure most calibration effects absolutely (as if in the laboratory) for use in calibration. The effects are usually far too changeable. Instead, the calibration is achieved by making observations of calibrator sources on the appropriate timescales for the relevant effects, and solving the measurement equation for them using the fact that we have $N_{ant}(N_{ant} - 1)/2$ measurements and only N_{ant} factors to determine (except for M which is only sparingly used). (*Note: By partitioning the calibration factors into a series of consecutive effects, it might appear that the number of free parameters is some multiple of N_{ant} , but the relative algebra and timescales of the different effects, as well as the multiplicity of observed polarizations and channels compensate, and it can be shown that the problem remains well-determined until, perhaps, the effects are direction-dependent within the field of view. Limited solvers for such effects are under study; the `calibrator` tool currently only handles effects which may be assumed constant within the field of view. Corrections for the primary beam are handled in the `imager` tool.*) Once determined, these terms are used to correct the visibilities measured for the scientific target. This procedure is known as cross-calibration (when only phase is considered, it is called phase-referencing).

The best calibrators are point sources at the phase center (constant visibility amplitude, zero phase), with sufficient flux density to determine the calibration factors with adequate SNR on the relevant timescale. The primary gain calibrator must be sufficiently close to the target on the sky so that its observations sample the same atmospheric effects. A bandpass calibrator usually must be sufficiently strong (or observed with sufficient duration) to provide adequate per-channel sensitivity for a useful calibration. In practice, several calibrators are usually observed, each with properties suitable for one or more of the required calibrations.

Synthesis calibration is inherently a bootstrapping process. First, the dominant calibration term is determined, and then, using this result, more subtle effects are solved for, until the full set of required calibration terms is available for application to the target field. The solutions for each successive term are relative to the previous terms. Occasionally, when the several calibration terms are not sufficiently orthogonal, it is useful to re-solve for earlier types using the results for later types, in effect, reducing the effect of the later terms on the solution for earlier ones, and thus better isolating them. This idea is a generalization of the traditional concept of self-calibration, where initial imaging of the target source supplies the visibility model for a re-solve of the gain calibration (G or T). Iteration tends toward convergence to a statistically optimal image. In general, the quality of each calibration and of the source model are mutually dependent. In principle, as long as the solution for any calibration component (or the source model itself) is likely to improve substantially through the use of new information (provided by other improved solutions), it is worthwhile to continue this process.

In practice, these concepts motivate certain patterns of calibration for different types of observation, and the `calibrator` tool in CASA is designed to accommodate these patterns in a general and flexible manner. For a spectral line total intensity observation, the pattern is usually:

1. Solve for G on the bandpass calibrator
2. Solve for B on the bandpass calibrator, using G

3. Solve for G on the primary gain (near-target) and flux density calibrators, using B solutions just obtained
4. Scale G solutions for the primary gain calibrator according to the flux density calibrator solutions
5. Apply G and B solutions to the target data
6. Image the calibrated target data

If opacity and gain curve information are relevant and available, these types are incorporated in each of the steps (in future, an actual solve for opacity from appropriate data may be folded into this process):

1. Solve for G on the bandpass calibrator, using T (opacity) and E (gain curve) solutions already derived.
2. Solve for B on the bandpass calibrator, using G , T (opacity), and E (gain curve) solutions.
3. Solve for G on primary gain (near-target) and flux density calibrators, using B , T (opacity), and E (gain curve) solutions.
4. Scale G solutions for the primary gain calibrator according to the flux density calibrator solutions
5. Apply T (opacity), E (gain curve), G , and B solutions to the target data
6. Image the calibrated target data

For continuum polarimetry, the typical pattern is:

1. Solve for G on the polarization calibrator, using (analytical) P solutions.
2. Solve for D on the polarization calibrator, using P and G solutions.
3. Solve for G on primary gain and flux density calibrators, using P and D solutions.
4. Scale G solutions for the primary gain calibrator according to the flux density calibrator solutions.
5. Apply P , D , and G solutions to target data.
6. Image the calibrated target data.

For a spectro-polarimetry observation, these two examples would be folded together.

In all cases the calibrator model must be adequate at each solve step. At high dynamic range and/or high resolution, many calibrators which are nominally assumed to be point sources become slightly resolved. If this has biased the calibration solutions, the offending calibrator may be imaged at any point in the process and the resulting model used to improve the calibration. Finally, if sufficiently strong, the target may be self-calibrated as well.

E.2 General Calibrator Mechanics

The calibrator tasks/tool are designed to solve and apply solutions for all of the solution types listed above (and more are in the works). This leads to a single basic sequence of execution for all solves, regardless of type:

1. Set the calibrator model visibilities
2. Select the visibility data which will be used to solve for a calibration type
3. Arrange to apply any already-known calibration types (the first time through, none may yet be available)
4. Arrange to solve for a specific calibration type, including specification of the solution timescale and other specifics
5. Execute the solve process
6. Repeat 1-4 for all required types, using each result, as it becomes available, in step 2, and perhaps repeating for some types to improve the solutions

By itself, this sequence doesn't guarantee success; the data provided for the solve must have sufficient SNR on the appropriate timescale, and must provide sufficient leverage for the solution (e.g., D solutions require data taken over a sufficient range of parallactic angle in order to separate the source polarization contribution from the instrumental polarization).

Appendix F

Data Weights

F.1 Introduction

Visibility weight initialization and calibration has undergone several improvements in CASA 4.2.2 and (pending) CASA 4.3. This appendix briefly describes the formal weight definitions, and the changes occurring in these CASA versions. If data sets shall be combined that were reduced with different CASA versions, the weights may need to be adjusted accordingly. This can be achieved, e.g. by running the same version of `statwt` (§ 4.7.9) on all datasets before combination. The best option, however, is to use a single CASA version for all reductions, preferably 4.2.2 or later.

Note that post-calibration weights, e.g. imaging weights or tapers are not covered by this appendix.

F.2 SIGMA and WEIGHT columns

Formally, in CASA 4.2.2 and later, the `SIGMA` column in the measurement set will reflect the per-channel noise of the `DATA` as it depends on the channel bandwidth $\Delta\nu$ and the length of an integration Δt :

$$\text{SIGMA} = \frac{1}{\sqrt{2\Delta\nu\Delta t}}. \tag{F.1}$$

The factor of $\sqrt{2}$ is for cross-correlation only and auto-correlation data follows $\text{SIGMA} = 1/\sqrt{\Delta\nu\Delta t}$.

`SIGMA` will only be updated if the time and channel widths are modified along with any `DATA` column manipulation, e.g. through averaging, binning, smoothing, etc. (tasks like `mstransform`, `cvel`, `split`, `exportuvfits`,...).

The `WEIGHT` column reflects how much weight each `CORRECTED_DATA` sample should receive when data are combined (e.g., in averaging). To start with, `WEIGHT` is initialized from the `SIGMA` column via:

	\leq CASA 4.2.1	CASA 4.2.2/4.3	\geq CASA 4.4 (WEIGHT_SPECTRUM)
Initialization	1	$2\Delta\nu\Delta t$	$2\Delta\nu\Delta t$
System Temperature	$\frac{1}{\langle\sqrt{T_{\text{sys},k}}\rangle_k^2}$	$\frac{1}{\langle T_{\text{sys},k}\rangle_k}$	$\frac{1}{T_{\text{sys},k}}$
Gains	$\ G\ ^2$	$\ G\ ^2$	$\ G\ ^2$
Bandpass	$\frac{1}{\langle\ B\ ^{-1}\rangle_k^2}$	$\langle\ B\ ^2\rangle_k$	$\langle\ B\ ^2\rangle_k$

Table F.1: Antenna-based WEIGHT calibration factor definitions for different CASA versions. For System Temperature and Bandpass, k is the channel index. ALMA has channelized T_{sys} ; EVLA does not.

$$\text{WEIGHT} = \frac{1}{\text{SIGMA}^2} = 2\Delta\nu\Delta t \quad (\text{F.2})$$

Data calibration by `applycal` (§ 4.6.1) with `calwt=T` will calculate and modify the WEIGHT values but not SIGMA. Calibration applies multiplicative factors and the WEIGHT of a visibility on a baseline between antennas i and j is calculated via

$$\text{WEIGHT}_{ij} = \frac{\omega_i\omega_j}{\text{SIGMA}_{ij}^2} \quad (\text{F.3})$$

where ω_i and ω_j are the antenna-based calibration factors derived by `applycal` ($\omega_i = \omega_j$ for auto-correlation data). In Table F.1 we list the definitions of *antenna-based* ω for different calibration procedures and CASA versions. When more than one calibration is applied, the product of the relevant weight factors is used.

F.2.1 Weights in CASA 4.2.1 and Earlier

The SIGMA and WEIGHT columns are initialized with values of “1”. Traditionally, this convention was adequate for datasets with uniform sampling in time in frequency; a global weight scale factor would not affect calibration and imaging fidelity. In data manipulation operations (e.g., `split`, etc.), SIGMA was treated as a per-channel value and WEIGHT as a per-spw (all channels) weight. Combined with unit initialization, this difference in definition could lead to incongruent weight scales for different spectral windows, in particular if bandwidth and channel count varied. CASA 4.2.1 is *not* recommended for datasets which have variety in spectral window bandwidth and channelization and for which spectral windows are to be combined in imaging.

F.2.2 Weights in CASA 4.2.2

In CASA 4.2.2 the `SIGMA` and `WEIGHT` columns are properly initialized via the definition in Eqs. F.1 and F.2. Both are defined as per-channel values. Also, the weight calibration factors have been subtly updated to improve robustness, as indicated in Table F.1.

F.2.3 Weights in CASA 4.3

In CASA 4.3 frequency variations of the `WEIGHT` and `SIGMA` values are (optionally) captured in additional `WEIGHT_SPECTRUM` and `SIGMA_SPECTRUM` columns. This allows accommodation of variations of effective sensitivity on a channel by channel basis (e.g. band edges, atmospheric lines, spectral T_{sys} variation etc.). `WEIGHT_SPECTRUM` will be recognized in the `applycal` task as well as in `mstransform` and `clean`. Calibration solvers, however, will not yet calculate and modify `WEIGHT_SPECTRUM`.

F.2.4 Weights in CASA 4.4 and later

Full support of `WEIGHT_SPECTRUM`. Also, `statwt` will be enhanced to permit calculating weights on a channel-dependent basis.

Appendix G

Cal Library Syntax

G.1 Introduction

The "Cal Library" is a new means of expressing calibration application instructions. It has nominally been available in `applycal` and the calibration solve tasks since CASA 4.1, via the `docallib=True` parameter, as an alternative to the traditional parameters (e.g., `gaintable`, etc.) that most users continue to use. As of CASA 4.5, we have deployed use of the Cal Library for *on-the-fly* calibration in `plotms` and `mstransform`. In CASA 4.5, our intent is to demonstrate the Cal Library and begin familiarizing users with it. The capabilities remain limited in some ways, and new features, additional flexibility, and broader deployment in more tasks will be offered in CASA 4.6 and later releases.

This appendix describes basic use of the Cal Library.

Alert: Please note the section on current (CASA 4.5) limitations.

G.2 Basic Cal Library Usage

The Cal Library is a means of specifying calibration instructions in an ascii file, rather than via the traditional `gaintable/gainfield/interp/spwmap/calwt` parameters that often become clumsy when many caltables are involved, and which have rather limited flexibility. Instead of specifying the traditional parameters, the file name is specified in the `callib` parameter in `applycal` or `plotms` (in `applycal` one must also specify `docallib=T`). For example, to correct an MS called *my.ms*, with a Cal Library file called *mycal.txt*:

```
applycal(vis='my.ms',docallib=T,callib='mycal.txt')
```

In a Cal Library file, each row expresses the calibration apply instructions for a particular caltable and (optionally) a specific selection of data in the MS to which it is to be applied.

For example, if *mycal.txt* contains:

```
caltable='cal.G' tinterp='linear' calwt=True
```

this will arrange a caltable called *cal.G* to be applied (with no detailed selection) to all MS data with linear interpolation in time, and with the weights also calibrated. It corresponds to these settings for the traditional parameters in applycal:

```
applycal(vis='my.ms',gaintable='cal.G',gainfield='',interp='linear',
         spwmap=[],calwt=True)
```

If a bandpass table, *cal.B*, is also available for application, one might use the following Cal Library file:

```
caltable='cal.G' tinterp='linear' calwt=True
caltable='cal.B' finterp='linear' calwt=False
```

This example arranges the same instructions for *cal.G*, and adds a bandpass table that will be interpolated linearly in frequency (the default for time-dependent interpolation is linear, if the bandpass table contains more than one time sample), without weight calibration. The corresponding form with the traditional parameters is:

```
applycal(vis='my.ms',gaintable=['cal.G','cal.B'], gainfield=['',''],
         interp=['linear','linear,linear'],
         spwmap=[],calwt=[True,False])
```

In general, the Cal Library file should be easier to read and manage than the traditional parameters as the number of specified caltables grows.

A more complicated example, involving non-trivial *spwmap* as well as field selection (*fldmap*) in the caltable:

```
caltable='cal.G' tinterp='linear' fldmap='nearest' spwmap=[0,1,1,3] calwt=True
caltable='cal.B' finterp='linear' fldmap='3' spwmap=[0,0,0,0] calwt=False
```

In this case, solutions from *cal.G* will be selected based on directional proximity ('nearest') for each MS field via the *fldmap* parameter, and spw 2 will be calibrated by spw 1 solutions. For *cal.B*, solutions from field id 3 will be used exclusively, with spw 0 calibrating all MS spws (of which there are apparently 4). The corresponding settings for the traditional parameters is as follows:

```
applycal(vis='my.ms',gaintable=['cal.G','cal.B'], gainfield=['nearest','3'],
         interp=['linear','linear,linear'],
         spwmap=[[0,1,1,3],[0,0,0,0]],calwt=[True,False])
```

Comment lines may be included in the cal library file by starting a line with the # character. (Partial line comments are *not* supported, as yet.) Existing cal library lines can be turned off (for experimentation purposes by making those lines comments with #.

G.3 More Advanced Cal Library Usage

The real power of the Cal Library arises from the ability to specify calibration instructions for a caltable *per MS selection*. This enables consolidating what would be multiple applycal executions using the traditional parameters into a single execution. Extending the example from above, if the MS field 'cal' should be calibrated by *cal.G* with 'nearest' interpolation in time, and the field 'sci' with 'linear' interpolation in time, the following Cal Library file will achieve this:

```
caltable='cal.G' field='cal' tinterp='nearest' fldmap='nearest' spwmap=[0,1,1,3] calwt=True
caltable='cal.G' field='sci' tinterp='linear' fldmap='nearest' spwmap=[0,1,1,3] calwt=True
caltable='cal.B' finterp='linear' fldmap='3' spwmap=[0,0,0,0] calwt=False
```

Note that the algorithm for selecting solutions from the caltable (`fldmap='nearest'`, which may resolve differently for the two MS fields) hasn't been changed, but it could be. In fact, any of the calibration parameters can be adjusted per MS selection, except `calwt`, which if set to `True` for any MS selection, will be forced to `True` for all (to maintain weight consistency within the MS). MS selection by `spw`, `intent`, and `obs id` can also be used (see the glossary below).

The pair of applycal executions corresponding to this Cal Library would be:

```
applycal(vis='my.ms',field='cal',gaintable=['cal.G','cal.B'], gainfield=['nearest','3'],
         interp=['nearest','linear,linear'], spwmap=[[0,1,1,3],[0,0,0,0]],calwt=[True,False])

applycal(vis='my.ms',field='sci',gaintable=['cal.G','cal.B'], gainfield=['nearest','3'],
         interp=['linear','linear,linear'], spwmap=[[0,1,1,3],[0,0,0,0]],calwt=[True,False])
```

G.4 General Rules (current, as of CASA 4.5)

- Each non-comment line in the Cal Library file must contain a valid (existing) caltable name.
- Blank lines (i.e., containing whitespace only) will be ignored.
- All parameters (see glossary below) are name/value pairs using an equals sign, delimited with spaces (no commas!)
- Only those parameters (see glossary) for which non-default values are required need be specified
- Each set of coordinated instructions must occur on a single line (there is no line continuation operator, as yet)
- If detailed MS selection is used, care must be exercised to ensure it is mutually exclusive over all MS rows for the same caltable; there is currently no internal checking for redundancy, and only the last calibration instructions for a particular MS selection will be invoked
- Full-line comments are supported by inserting the # character as the first non-whitespace character in the line. This mechanism can be used to turn off ordinary cal library lines.

G.5 Limitations (current, as of CASA 4.5)

- Application of parallactic angle corrections is not yet supported within the Cal Library file (this only affects use in `plotms`, where there is no `parang` parameter)
- Some parametrized calibration tables (`BPOLY`, `GSPLINE`) are not yet supported

G.6 Conversion from Existing `applycal` Scripts

To convert exiting `applycal` commands, a simple experimental function, `applycaltocallib` is available. To access it, type (within CASA):

```
from callibrary import applycaltocallib
```

Then, chose a filename for the cal library file, and supply existing settings for `applycal` parameters (`field`, `spw`, `intent`, `gaintable`, `gainfield`, `interp`, `spwmap`, `calwt`) to the `applycaltocallib` function:

```
callibfile='mycallib.txt'
applycaltocallib(filename=callibfile,append=F,
                 field,spw,intent,gaintable,gainfield,
                 interp,spwmap,calwt)
```

If `append=F`, the specified `filename` will be overwritten, if it already exists. If `append=T`, new entries will be appended to the existing `filename`. Only parameters with non-trivial `applycal` settings need be included. In general, if `gaintable` is a python list, it is best if `gainfield`, `interp`, `spwmap`, and `calwt` (where non-trivially set) are also lists.

For example, if your conventional script contains the following `applycal` executions (duplicated from above):

```
applycal(vis='my.ms',field='cal',
         gaintable=['cal.G','cal.B'], gainfield=['nearest','3'],
         interp=['nearest','linear,linear'],
         spwmap=[[0,1,1,3],[0,0,0,0]],
         calwt=[True,False])
applycal(vis='my.ms',field='sci',
         gaintable=['cal.G','cal.B'], gainfield=['nearest','3'],
         interp=['linear','linear,linear'],
         spwmap=[[0,1,1,3],[0,0,0,0]],
         calwt=[True,False])
```

...these can be edited to `applycaltocallib` executions as:


```

callibfile='mycallib.txt'
applycaltocallib(filename='mycallib.txt',append=F,
                 field='cal',
                 gaintable=['cal.G','cal.B'], gainfield=['nearest','3'],
                 interp=['nearest','linear,linear'],
                 spwmap=[[0,1,1,3],[0,0,0,0]],
                 calwt=[True,False])
applycaltocallib(filename='mycallib.txt',append=T,
                 field='sci',
                 gaintable=['cal.G','cal.B'],
gainfield=['nearest','3'],
                 interp=['linear','linear,linear'],
                 spwmap=[[0,1,1,3],[0,0,0,0]],
                 calwt=[True,False])

```

After running them, `mycallib.txt` will contain:

```

caltable='cal.B' calwt=False field='cal' tinterp='linear' finterp='linear' fldmap='3' spwmap=[0, 0, 0, 0]
caltable='cal.G' calwt=True field='cal' tinterp='nearest' fldmap='nearest' spwmap=[0, 1, 1, 3]
caltable='cal.B' calwt=False field='sci' tinterp='linear' finterp='linear' fldmap='3' spwmap=[0, 0, 0, 0]
caltable='cal.G' calwt=True field='sci' tinterp='linear' fldmap='nearest' spwmap=[0, 1, 1, 3]

```

Note that the `cal.B` table is specified separately for the `'cal'` and `'sci'` fields with otherwise the same parameters; thus, those two lines could be manually consolidated to a single line with unified field selection, yielding:

```

caltable='cal.B' calwt=False field='cal,sci' tinterp='linear' finterp='linear' fldmap='3' spwmap=[0, 0, 0, 0]
caltable='cal.G' calwt=True field='cal' tinterp='nearest' fldmap='nearest' spwmap=[0, 1, 1, 3]
caltable='cal.G' calwt=True field='sci' tinterp='linear' fldmap='nearest' spwmap=[0, 1, 1, 3]

```

The field selection for the first row could be removed entirely if `cal.B` will be used uniformly for all fields in the MS (equivalently, `field=''`). This sort of row consolidation is optional, but it may have useful memory efficiency benefits when running `applycal`, and so is recommended.

The `applycaltocallib` function should be considered experimental and used with care, and the resulting file examined thoroughly for correctness, since this function will not do any internal duplication checking or other sanity checks. All other current constraints and limitations on cal libraries (as noted above) will apply.

G.7 Glossary

This is a list of recognized Cal Library parameters. For each, the default is indicated. Additional parameters enhancing flexibility will be added in CASA 4.5 and later.

- `caltable` — the name of the caltable for which the instructions on the current line apply; no default; required

G.7.1 MS selection

Use these parameters to implement calibration instructions specific to particular MS selections (using standard MS Selection syntax, except where noted).

- **field** — the MS field selection for which the calibration instructions on the current line apply; default='' (all fields)
- **spw** — the MS spw selection for which the calibration instructions on the current line apply; default='' (all spws) Note that channel selection will be ignored, since the Cal Library does not support variety in calibration application at channel granularity.
- **intent** — the MS intent selection for which the calibration instructions on the current line apply; default='' (all intents)
- **obs** — the MS observation id selection for which the calibration instructions on the current line apply; default='' (all intents)

G.7.2 Interpolation/application

- **tinterp** — the time-dependent interpolation mode; default='linear' options: 'linear', 'nearest'
- **finterp** — the chan-dependent interpolation mode (only relevant for channelized caltables); default='linear' options: 'nearest', 'linear', 'cubic', 'spline'
- **calwt** — weight calibration; default=True options: True, False

G.7.3 Calibration Mapping

The following ***map** parameters enable selection on the caltable. For each ***map** parameter, the basic specification is an ordered list indicating the caltable selection indices intended for each MS index on that axis. E.g., **spwmap**=[0,1,1,3] means MS spws 0,1,3 will each be calibrated by the same spw index from the caltable, and MS spw 2 will be calibrated by cal spw 1. The ***map** parameters support other short-hand options as well, as indicated below. For defaults, “index identity” means that each MS index will be calibrated by the corresponding caltable index, and “no explicit mapping” means that no filter will be applied to that axis, and all available solutions on the axis will be included.

- **spwmap** — spectral window mapping; default=index identity
- **fldmap** — field mapping; default=[] (no explicit mapping); additional options: 'nearest' or a string indicating selection on the caltable (same as traditional **gainfield** options)
- **antmap** — antenna id mapping; default=index identity
- **obsmap** — obs id mapping; default=[] (no explicit mapping)

Appendix H

Annotated Example Scripts

<http://casaguides.nrao.edu> is the source for a variety of data reduction examples for ALMA, the VLA, simulations, and other data.

Appendix I

CASA Dictionaries

BETA ALERT: These tend to become out of date as we add new tasks or change names.

I.1 AIPS – CASA dictionary

In Table I.1 we provide a comparison of CASA and AIPS commands. The data reduction recipes and break-down of jobs in individual tasks, however, is not the same in both packages. Nevertheless, the table may give AIPS users a good start if they search for functionality in CASA.

I.2 MIRIAD – CASA dictionary

Table I.2 provides a list of common Miriad tasks, and their equivalent CASA tool or tool function names. The two packages differ in both their architecture and calibration and imaging models, and there is often not a direct correspondence. However, this index does provide a scientific user of CASA who is familiar with MIRIAD, with a simple translation table to map their existing data reduction knowledge to the new package.

I.3 CLIC – CASA dictionary

Table I.3 provides a list of common CLIC tasks, and their equivalent CASA tool or tool function names. The two packages are very similar since the CASA software to reduce IRAM data is based on the CLIC reduction procedures.

Table I.1: AIPS – CASA dictionary

AIPS Task	CASA task/tool	Description
APROPOS	taskhelp	List tasks with a short description of their purposes
BLCAL	blcal	Calculate a baseline-based gain calibration solution
BLCHN	blcal	Calculate a baseline-based bandpass calibration solution
BPASS	bandpass	Calibrate bandpasses
CALIB	gaincal	Calibrate gains (amplitudes and phases)
CLCAL	applycal	Apply calibration to data
COMB	immath	Combine images
CPASS	bandpass	Calibrate bandpasses by polynomial fitting
DBCON	concat	Concatenate u-v datasets
DEFAULT	default	Load a task with default parameters
FILLM	importvla	Import old-format VLA data
FITLD	importuvfits	Import a u-v dataset which is in FITS format
FITLD	importfits	Import an image which is in FITS format
FITTP	exportuvfits	Write a u-v dataset to FITS format
FITTP	exportfits	Write an image to FITS format
FRING	—	Calibrate group delays and phase rates.
GETJY	fluxscale	Determine flux densities for other cal
GO	go	Run a task
HELP	help	Display the help page for a task
IMAGR	clean	Image and deconvolve
IMFIT	imfit	Fit gaussian components to an image
IMHEAD	vishead	View header for u-v data
IMHEAD	imhead	View header for an image
IMLIN	imcontsub	Subtract continuum in image plane
IMLOD	importfits	Import a FITS image
IMSTAT	imstat	Measure statistics on an image
INP	inp	View task parameters
JMFIT	imfit	Fit gaussian components to an image
LISTR	listobs	Print basic data
MCAT	ls	List image data files
MOMNT	immoments	Compute moments from an image
OHGEO	imregrid	Regrids an image onto another image's geometry
PBCOR	immath	Correct an image for the primary beam
PCAL	polcal	Calibrate polarization
POSSM	plotcal	Plot bandpass calibration tables
POSSM	plotms	Plot spectra
PRTAN	listobs	Print antenna locations
PRTAN	plotants	Plot antenna locations
QUACK	flagdata	Remove first integrations from scans
RENAME	mv	Rename an image or dataset
SETJY	setjy	Set flux densities for flux cal
SMOTH	imsmooth	Smooth an image
SNPLT	plotcal	Plot gain calibration tables
SPFLG	viewer	Flag raster image of time v. channel
SPLIT	split	Write out u-v files for individual sources
TASK	inp	Load a task with current parameters
TGET	tget	Load a task with parameters last used for that task
TVALL	viewer	Display image
TVFLG	viewer	Flag raster image of time v. baseline
UCAT	ls	List u-v data files
UVFIX	fixvis	Compute u, v, and w coordinates
UVFLG	flagdata	Flag data
UVLIN	uvcontsub	Subtract continuum from u-v data
UVLSF	uvcontsub	Subtract continuum from u-v data
UVPLT	plotms	Plot u-v data
UVSUB	uvsb	Subtracts model u-v data from corrected u-v data
WIPER	plotms	Plot and flag u-v data
ZAP	rmtables	Delete data files

Table I.2: MIRIAD – CASA dictionary

MIRIAD Task	Description	CASA task/tool
blflag	Interactive baseline based editor/flagger	mp raster displays
cg curs	Interactive image analysis	viewer
cgdisp	Image display, overlays	viewer
clean	Clean an image	clean
fits	FITS image filler	importfits, exportfits, importuvfits, exportuvfits
gpboot	Set flux density scale	fluxscale
gpcal	Polarization leakage and gain calibration	gaincal
gpcopy	copy calibration tables	<i>not needed</i>
gpplt	Plot calibration solutions	plotcal
imcomb	Image combination	immaths
imfit	Image-plane component fitter	imfit
impol	Create polarization images	clean
imstat	Image statistics	imstats
imsub	Extract sub-image	ia.subimage
invert	Synthesis imaging	clean
linmos	linear mosaic combination of images	clean
maths	Calculations involving images	immath
mfc al	Bandpass and gain calibration	bandpass
prthd	Print header of image or uvdata	imhead, listobs, vishead
restor	Restore a clean component model	clean
selfcal	selfcalibration of visibility data	clean, gaincal, etc.

Table I.3: CLIC–CASA dictionary

CLIC Function	Description	CASA task/tool
load	Load data	importfits, importasdm, importuvfits
print	Print text summary of data	listobs
flag	Flag data	plotms, flagdata, viewer
phcor	Atmospheric phase correction	gaincal
rf	Radio frequency bandpass	bandpass
phase	Phase calibration	gaincal
flux	Absolute flux calibration	setjy, fluxscale
ampl	Amplitude calibration	gaincal
table	Split out calibrated data (uv table)	split

Appendix J

Writing Tasks

ALERT: This prescription for writing and incorporating tasks in CASA is for the power-user. This procedure is also likely to change in future releases.

It is possible to write your own task and have it appear in casapy. For example, if you want to create a task named “yourtask”, then must create two files, `yourtask.xml` and a `task_yourtask.py`. The `.xml` file is use to describe the interface to the task and the `task_yourtask.py` does the actual work. The argument names must be the same in both the `yourtask.xml` and `task_yourtask.py` file. The `yourtask.xml` file is used to generate all the interface files so `yourtask` will appear in the casapy system. It is easiest to start from one of the existing tasks when constructing these. You would make the name of the function in the `yourtask.py` be “yourtask” in this example.

We have provided the `buildmytasks` command in order to assemble your Python and XML into a loadable Python file. Thus, the steps you need to execute (again for an example task named “yourtask”):

- Create python code for task as `task_yourtask.py`
- Create xml for task as `yourtask.xml`
- Execute `buildmytasks` from the casapy prompt: `!buildmytasks`
- Initialize your new task inside casapy: `execfile 'mytasks.py'`

After this, you should see the help and inputs inside casapy, e.g. `inp yourtask` should work. Note that for the final step you invoke the file called `mytasks.py`, regardless of what you named the actual task. You now have a shiny new task `yourtask` that you can run and use in the same way as all other CASA tasks.

Note that if multiple custom tasks are stored in the same directory, they will all be built by `!buildmytasks` and will all be initialized by executing `mytasks.py`. To build and initialize only a single task, instead use `!buildmytasks taskname`; you are then free to rename `mytasks.py` (e.g. `load_taskname.py`) and repeat this procedure for your other tasks. Our recommendation, for those of you who are managing multiple custom tasks, is to have each task live in its own directory. The

`mytasks.py` file need not be in the current working directory to initialize your task, since you can provide the full path upon initialization (e.g. `execfile '/full_path_to_my_task/mytasks.py'`).

J.1 The XML file

The key to getting your task into `casapy` is constructing a task interface description XML file.

Some XML basics, an xml element begins with `<element>` and ends with `</element>`. If an XML element contains no other XML element you may specify it via `<element/>`. An XML element may have zero or more attributes which are specified by `attribute="attribute value"`. You must put the attribute value in quotes, i.e. `<element myattribute="attribute value">`.

All task xml files must start with this header information.

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" ?>
<casaxml xmlns="http://casa.nrao.edu/schema/psetTypes.html"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://casa.nrao.edu/schema/casa.xsd
file:///opt/casa/code/xmlcasa/xml/casa.xsd">
```

and the file must have the end tag

```
</casaxml>
```

Inside a `<task>` tags you will need to specify the following elements.

```
<task>
```

Attributes

- type required, allowed value is "function"
- name required

Subelements

- shortdescription** required
- description** required
- input** optional
- output** optional
- returns** optional
- constraints** optional

`<shortdescription>` - required by `<task>`; A short one-line description describing your task

Attributes

None

Subelements

None

<**description**> - required] by <task>, Also used by <param>a; A longer description describing your task with multiple lines

Attributes

None

Subelements

None

<**input**> - optional element used by <task>; An input block specifies which parameters are used for input

Attributes

None

Subelements

<**param**> , optional

<**output**> - **optional** An output element that contains a list of parameters that are "returned" by the task.

Attributes

None

Subelements

<**param**> , optional

<**returns**> - **optional** Value returned by the task

Attributes

type optional; as specified in <param>

Subelements

<**description**> , optional

<**constraints**> - **optional** A constraints element that lets you constrain params based on the values of other params.

Attributes

None

Subelements

<**when**> , required.

<**param**> - **optional** The input and output elements consist of param elements.

Attributes

type , required; allowed values are record, variant, string int, double, bool, intArray, doubleArray, boolArray, stringArray

name , required;

subparam , optional; allowed values True, False, Yes or No.

kind , optional;

mustexist , optional; allowed values True, False, Yes or No.

All param elements require name and type attributes.

Subelements

<description> , required;

<value> , optional;

<allowed> , optional;

<value> - **optional** Value returned by the task

Attributes

type , required; as specified in <param> attributes.

Subelements

<value> , optional

<allowed> - **optional**; Block of allowed values

Attributes

enum , required; maybe enum or range. If specified as enum only specific values are allowed If specified as range then the value tags may have min and max attributes.

Subelements

<value> , optional

<when> - **optional** When blocks allow value specific handling for parameters

Attributes

param , required; Specifies special handling for a <param>

Subelements

<equals> , optional

<notequals> , optional

<equals> - **optional** Reset parameters if equal to the specified value

Attributes

value , required; the value of the parameter

Subelements

<default> , required

<notequals> - **optional** Reset specified parameters if not equal to the specified value

Attributes

value , required; The value of the parameter

Subelements

<default> , optional

<default> - **optional** Resets default values for specified parameters

Attributes

param , required; Name of the **<param>** to be reset.

Subelements

<value> , required, the revised value of the **<param>**.

<example> - **optional** An example block, typically in python

Attributes

lang optional; specifies the language of the example, defaults to python.

Subelements

None

J.2 The `task_yourtask.py` file

You must write the python code that does the actual work. The `task_*.py` file function call sequence must be the same as specified in the XML file. We may relax the requirement that the function call sequence exactly match the sequence in the XML file in a future release.

The `task_*.py` file should contain the following preamble

```
import os
from taskinit import *
```

plus any other global function imports you will need such as

```
import time
```

followed by the task function `def`. See Sect. J.3.2 for an example.

J.3 Example: The clean task

Note that the following is for illustration only and does *not* reflect the current implementation of `clean` – a task that is always developing further on a quick pace.

J.3.1 File clean.xml

Clean.xml gives a fairly comprehensive example of how to construct the XML file.

```
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" ?>
<casaxml xmlns="http://casa.nrao.edu/schema/psetTypes.html"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://casa.nrao.edu/schema/casa.xsd
file:///opt/casa/code/xmlcasa/xml/casa.xsd">

<!-- This is the param set for clean -->
<!-- This does the equivalent of -->
<!-- imgr:=imager('anyfile.ms'); -->
<!-- imgr.setdata(mode='channel',nchan=100,start=1,step=1,fieldid=1) -->
<!-- imgr.setimage(nx=512,ny=,cellx='1arcsec',celly='1arcsec',stokes='I',-->
<!--           mode='channel',start=35,step=1,nchan=40, -->
<!--           fieldid=[1]) -->
<!-- imgr.weight('natural'); -->
<!-- imgr.clean(algorithm='csclean',niter=500,model='field1') -->

<task type="function" name="clean">

  <shortdescription>Deconvolve an image with selected algorithm</shortdescription>

  <description>
  Form images from visibilities. Handles continuum and spectral line cubes.
  </description>

  <input>

    <param type="string" name="vis" kind="ms" mustexist="true">
    <description>name of input visibility file</description>
    <value></value>
    </param>

    <param type="string" name="imagenam">
      <description>Pre-name of output images</description>
      <value></value>
    </param>

    <param type="string" name="field">
      <description>Field Name</description>
```

```

    <value></value>
</param>

<param type="any" name="spw">
  <description>Spectral windows:channels: '\'\ ' is all </description>
  <any type="variant"/>
  <value type="string"></value>
</param>
<param type="bool" name="selectdata">
  <description>Other data selection parameters</description>
  <value>False</value>
</param>

<param type="string" name="timerange" subparam="true">
  <description>Range of time to select from data</description>
  <value></value>
</param>
<param type="string" name="uvrange" subparam="true">
  <description>Select data within uvrange </description>
  <value></value>
</param>
<param type="string" name="antenna" subparam="true">
  <description>Select data based on antenna/baseline</description>
  <value></value>
</param>
<param type="string" name="scan" subparam="true">
  <description>scan number range</description>
  <value></value>
</param>

<param type="string" name="mode">
  <description>
    Type of selection (mfs, channel, velocity, frequency)
  </description>
  <value>mfs</value>
  <allowed kind="enum">
    <value>mfs</value>
    <value>channel</value>
    <value>velocity</value>
    <value>frequency</value>
  </allowed>
</param>

```

```

<param type="int" name="niter">
<description>Maximum number of iterations</description>
<value>500</value>
</param>

<param type="double" name="gain">
  <description>Loop gain for cleaning</description>
  <value>0.1</value>
</param>

<param type="double" name="threshold" units="mJy">
  <description>Flux level to stop cleaning. Must include units</description>
  <value>0.0</value>
</param>

<!-- Getting rid of this
  <param type="bool" name="csclean">
    <description>Use Cotton-Schwab style reconciliation with UV-data</description>
    <value>False</value>

  </param>
-->

<param type="string" name="psfmode">
  <description>method of PSF calculation to use during minor cycles</description>
  <value>clark</value>
  <allowed kind="enum">
    <value>clark</value>
    <value>hogbom</value>
  </allowed>
</param>

<param type="string" name="imagermode">
  <description> Use csclean or mosaic. If '\', use psfmode</description>
  <value></value>
  <allowed kind="enum">
    <value></value>
    <value>csclean</value>
    <value>mosaic</value>
  </allowed>

</param>

<param type="string" name="ftmachine" subparam="true">

```

```

    <description>Gridding method for the image</description>
    <value>mosaic</value>
    <allowed kind="enum">
      <value>mosaic</value>
      <value>ft</value>
      <value>sd</value>
      <value>both</value>
    </allowed>

</param>
<param type="bool" name="mosweight" subparam="true">
  <description>Individually weight the fields of the mosaic</description>
  <value>False</value>
</param>
<param type="string" name="scatype" subparam="true">
  <description>Controls scaling of pixels in the image plane.
    default='\SAULT\';
    example: scatype='\PBCOR\'. Options: '\PBCOR\','SAULT\'.</description>
  <value>SAULT</value>
  <allowed kind="enum">
    <value>SAULT</value>
    <value>PBCOR</value>
  </allowed>
</param>

<param type="intArray" name="multiscale">
  <description>set deconvolution scales (pixels),
    default: multiscale=[] (standard CLEAN)</description>
  <value type="vector">
    <value></value>
  </value>

</param>
<param type="int" name="negcomponent" subparam="true">
  <description>
    Stop cleaning if the largest scale finds this number of neg components
  </description>
  <value>0</value>
</param>

<param type="bool" name="interactive">
  <description>use interactive clean (with GUI viewer)</description>
  <value>False</value>

```

```

</param>

<param type="any" name="mask">
  <description>cleanbox(es), mask image(s), and/or region(s) used in cleaning
  </description>
  <any type="variant"/>
  <value type="stringArray"></value>
</param>

<param type="int" name="nchan" subparam="true">
  <description>Number of channels (planes) in output image</description>
  <value>1</value>
</param>

<param type="any" name="start" subparam="true">
  <description>First channel in input to use</description>
  <any type="variant"/>
  <value type="int">0</value>
</param>

<param type="any" name="width" subparam="true">
  <description>Number of input channels to average</description>
  <any type="variant"/>
  <value type="int">1</value>
</param>

<param type="intArray" name="imsize">
  <description>x and y image size in pixels, symmetric for single value
  </description>
  <value type="vector">
<value>256</value><value>256</value>
  </value>
</param>

<param type="doubleArray" name="cell" units="arcsec">
<description>x and y cell size. default unit arcsec</description>
<value type="vector"><value>1.0</value><value>1.0</value></value>
</param>

<param type="any" name="phasecenter">
  <description>Image phase center: position or field index</description>
  <any type="variant"/>
  <value type="string"></value>

```



```

</param>

<param type="string" name="restfreq">
  <description>rest frequency to assign to image (see help)</description>
  <value></value>
</param>

<param type="string" name="stokes">
  <description>Stokes params to image (eg I,IV, QU,IQUV)</description>
  <value>I</value>
<allowed kind="enum">
  <value>I</value>
  <value>IV</value>
  <value>QU</value>
  <value>IQUV</value>
  <value>RR</value>
  <value>LL</value>
  <value>RRL</value>
  <value>XX</value>
  <value>YY</value>
  <value>XXYY</value>
</allowed>
</param>

<param type="string" name="weighting">
<description>Weighting to apply to visibilities</description>
<value>natural</value>
<allowed kind="enum">
  <value>natural</value>
  <value>uniform</value>
  <value>briggs</value>
  <value>briggsabs</value>
  <value>radial</value>
  <value>superuniform</value>
</allowed>
</param>

<param type="double" name="robust" subparam='true'>
  <description>Briggs robustness parameter</description>
  <value>0.0</value>

```

```

    <allowed kind="range">
    <value range="min">-2.0</value>
    <value range="max">2.0</value>
</allowed>
</param>

<param type="bool" name="uvtaper">
  <description>Apply additional uv tapering of visibilities.</description>
  <value>False</value>
</param>

<param type="stringArray" name="outertaper" subparam="true">
  <description>uv-taper on outer baselines in uv-plane</description>
  <value type="vector">
    <value></value>
  </value>
</param>

<param type="stringArray" name="innertaper" subparam="true">
  <description>uv-taper in center of uv-plane</description>
  <value>1.0</value>
</param>

<param type="string" name="modelimage">
  <description>Name of model image(s) to initialize cleaning</description>
  <value></value>
</param>

<param type="stringArray" name="restoringbeam">
  <description>Output Gaussian restoring beam for CLEAN image</description>
  <value></value>
</param>

<param type="bool" name="pbcor">
  <description>Output primary beam-corrected image</description>
  <value>False</value>
</param>

<param type="double" name="minpb">
  <description>Minimum PB level to use</description>
  <value>0.1</value>
</param>

<param type="any" name="noise" subparam='true'>
  <description>noise parameter for briggs abs mode weighting</description>
  <any type="variant"/>

```

```

        <value type="string">1.0Jy</value>
</param>

<param type="int" name="npixels" subparam='true'>
    <description>number of pixels for superuniform or briggs weighting
    </description>
    <value>0</value>
</param>

<param type="int" name="npercycle" subparam='true'>
    <description>Number of iterations before interactive prompt</description>
    <value>100</value>
</param>
<param type="double" name="cyclefactor" subparam='true'>
    <description>change depth in between of csclean cycle</description>
    <value>1.5</value>
</param>
<param type="int" name="cyclespeedup" subparam='true'>
    <description>Cycle threshold doubles in this number of iteration</description>
    <value>-1</value>
</param>

<constraints>
    <when param="selectdata">
        <equals type="bool" value="False"/>
        <equals type="bool" value="True">
            <default param="timerange"><value type="string"></value>
            </default>
            <default param="uvrange"><value type="string"></value>
            </default>
            <default param="antenna"><value type="string"></value>
            </default>
            <default param="scan"><value type="string"></value>
            </default>
        </equals>
    </when>
    <when param="multiscale">
        <notequals type="vector" value="[]" >
            <default param="negcomponent"><value>-1</value>
            </default>
        </notequals>
    </when>
</constraints>

```

```

    </notequals>
</when>
<when param="mode">
  <equals value="mfs"/>
  <equals value="channel">
    <default param="nchan"><value>1</value></default>
    <default param="start"><value>0</value>
    <description>first input channel to use</description>
  </default>
  <default param="width"><value>1</value></default>
</equals>
<equals value="velocity">
  <default param="nchan"><value>1</value></default>
  <default param="start"><value type="string">0.0km/s</value>
  <description>Velocity of first image channel: e.g. \'0.0km/s\'
  </description>
</default>
  <default param="width"><value type="string">1km/s</value>
  <description>image channel width in velocity units:
  e.g. \'-1.0km/s\'</description>
</default>
</equals>
<equals value="frequency">
  <default param="nchan"><value>1</value></default>
  <default param="start"><value type="string">1.4GHz</value>
  <description>Frequency of first image channel: e.g. \'1.4GHz\'
  </description>
</default>
  <default param="width"><value type="string">10kHz</value>
  <description>Image channel width in frequency units:
  e.g. \'1.0kHz\'</description>
</default>
</equals>
</when>

<when param="weighting">
  <equals value="natural"/>
  <equals value="uniform"/>
  <equals value="briggs">
    <default param="robust"><value>0.0</value></default>
    <default param="npixels"><value>0</value>
    <description>number of pixels to determine uv-cell size
    0=> field of view</description>
  </default>
</equals>

```

```

    <equals value="briggsabs">
      <default param="robust"><value>0.0</value></default>
      <default param="noise"><value type="string">1.0Jy</value></default>
      <default param="npixels"><value>0</value>
        <description>number of pixels to determine uv-cell size
          0=> field of view</description>
      </default>
    </equals>
  </when>
  <when param="superuniform">
    <default param="npixels"><value>0</value>
      <description>number of pixels to determine uv-cell size
        0=> +/-3pixels</description>
    </default>
  </equals>
</when>
<when param="uvtaper">
  <equals type="bool" value="False"/>
  <equals type="bool" value="True">
    <default param="outertaper"><value type="vector"></value></default>
    <default param="innertaper"><value type="vector"></value></default>
  </equals>
</when>
<when param="interactive">
  <equals type="bool" value="False"/>
  <equals type="bool" value="True">
    <default param="npercycle"><value>100</value></default>
  </equals>
</when>
<when param="imagermode">
  <equals value=""/>
  <equals value="csclean">
    <default param="cyclefactor"><value>1.5</value></default>
    <default param="cyclespeedup"><value>-1</value></default>
  </equals>
  <equals value="mosaic">
    <default param="mosweight"><value>False</value></default>
    <default param="ftmachine"><value type="string">mosaic</value>
      </default>
    <default param="scaletype"><value type="string">SAULT</value>
      </default>
    <default param="cyclefactor"><value>1.5</value></default>
    <default param="cyclespeedup"><value>-1</value></default>
  </equals>
</when>
<!--Get rid of that soon

```

```

    <when param="mosaicmode">
      <equals type="bool" value="False"/>
      <equals type="bool" value="True">
        <default param="mosweight"><value>False</value></default>
        <default param="ftmachine"><value type="string">mosaic</value></default>
        <default param="scaletype"><value type="string">SAULT</value></default>
      </equals>
    </when>
-->
  </constraints>

  </input>

<returns type="void"/>

<example>

```

The main clean deconvolution task. It contains many functions

- 1) Make 'dirty' image and 'dirty' beam (psf)
- 2) Multi-frequency-continuum images or spectral channel imaging
- 3) Full Stokes imaging
- 4) Mosaicking of several pointings
- 5) Multi-scale cleaning
- 6) Interactive clean boxing
- 7) Initial starting model

```

vis -- Name of input visibility file
      default: none; example: vis='ngc5921.ms'
imagename -- Pre-name of output images:
            default: none; example: imagename='m2'
            output images are:
              m2.image; cleaned and restored image
                With or without primary beam correction
              m2.psf; point-spread function (dirty beam)
              m2.flux; relative sky sensitivity over field
              m2.model; image of clean components
              m2.residual; image of residuals
              m2.interactive.mask; image containing clean regions
field -- Select fields in mosaic. Use field id(s) or field name(s).
        ['go listobs' to obtain the list id's or names]
        default: ''=all fields
        If field string is a non-negative integer, it is assumed to
          be a field index otherwise, it is assumed to be a

```

```

    field name
    field='0~2'; field ids 0,1,2
    field='0,4,5~7'; field ids 0,4,5,6,7
    field='3C286,3C295'; field named 3C286 and 3C295
    field = '3,4C*'; field id 3, all names starting with 4C
spw -- Select spectral window/channels
NOTE: This selects the data passed as the INPUT to mode
default: ''=all spectral windows and channels
    spw='0~2,4'; spectral windows 0,1,2,4 (all channels)
    spw='0:5~61'; spw 0, channels 5 to 61
    spw='&lt;2'; spectral windows less than 2 (i.e. 0,1)
    spw='0,10,3:3~45'; spw 0,10 all channels, spw 3,
        channels 3 to 45.
    spw='0~2:2~6'; spw 0,1,2 with channels 2 through 6 in each.
    spw='0:0~10;15~60'; spectral window 0 with channels
        0-10,15-60
    spw='0:0~10,1:20~30,2:1;2;3'; spw 0, channels 0-10,
        spw 1, channels 20-30, and spw 2, channels, 1,2 and 3
selectdata -- Other data selection parameters
    default: True
&gt;&gt;&gt; selectdata=True expandable parameters
See help par.selectdata for more on these
timerange -- Select data based on time range:
    default = '' (all); examples,
    timerange = 'YYYY/MM/DD/hh:mm:ss~YYYY/MM/DD/hh:mm:ss'
    Note: if YYYY/MM/DD is missing date defaults to first
        day in data set
    timerange='09:14:0~09:54:0' picks 40 min on first day
    timerange= '25:00:00~27:30:00' picks 1 hr to 3 hr
        30min on NEXT day
    timerange='09:44:00' pick data within one integration
        of time
    timerange='&gt;10:24:00' data after this time
uvrange -- Select data within uvrange (default units meters)
    default: '' (all); example:
    uvrange='0~1000klambda'; uvrange from 0-1000 kilo-lambda
    uvrange='&gt;4klambda';uvranges greater than 4 kilo lambda
antenna -- Select data based on antenna/baseline
    default: '' (all)
    If antenna string is a non-negative integer, it is
        assumed to be an antenna index, otherwise, it is
        considered an antenna name.
    antenna='5&6'; baseline between antenna index 5 and
        index 6.
    antenna='VA05&VA06'; baseline between VLA antenna 5

```

```

                                and 6.
                                antenna='5&6;7&8'; baselines 5-6 and 7-8
                                antenna='5'; all baselines with antenna index 5
                                antenna='05'; all baselines with antenna number 05
                                    (VLA old name)
                                antenna='5,6,9'; all baselines with antennas 5,6,9
                                    index numbers
scan -- Scan number range.
    default: '' (all)
    example: scan='1~5'
    Check 'go listobs' to insure the scan numbers are in
        order.
mode -- Frequency Specification:
    NOTE: See examples below:
    default: 'mfs'
        mode = 'mfs' means produce one image from all
            specified data.
        mode = 'channel'; Use with nchan, start, width to specify
            output image cube. See examples below
        mode = 'velocity', means channels are specified in
            velocity.
        mode = 'frequency', means channels are specified in
            frequency.
<>>> mode expandable parameters (for modes other than 'mfs')
    Start, width are given in units of channels, frequency
        or velocity as indicated by mode, but only channel
        is complete.
nchan -- Number of channels (planes) in output image
    default: 1; example: nchan=3
start -- Start input channel (relative-0)
    default=0; example: start=5
width -- Output channel width in units of the input
        channel width (>1 indicates channel averaging)
    default=1; example: width=4
examples:
    spw = '0,1'; mode = 'mfs'
        will produce one image made from all channels in spw
            0 and 1
    spw='0:5~28^2'; mode = 'mfs'
        will produce one image made with channels
            (5,7,9,...,25,27)
    spw = '0'; mode = 'channel': nchan=3; start=5; width=4
        will produce an image with 3 output planes
        plane 1 contains data from channels (5+6+7+8)
        plane 2 contains data from channels (9+10+11+12)

```



```

    plane 3 contains data from channels (13+14+15+16)
    spw = '0:0~63^3'; mode='channel'; nchan=21; start = 0;
        width = 1
    will produce an image with 20 output planes
    Plane 1 contains data from channel 0
    Plane 2 contains data from channel 2
    Plane 21 contains data from channel 61
    spw = '0:0~40^2'; mode = 'channel'; nchan = 3; start =
        5; width = 4
    will produce an image with three output planes
    plane 1 contains channels (5,7)
    plane 2 contains channels (13,15)
    plane 3 contains channels (21,23)
psfmode -- method of PSF calculation to use during minor cycles:
    default: 'clark': Options: 'clark','hogbom'
    'clark' use smaller beam (faster, usually good enough)
    'hogbom' full-width of image (slower, better for poor
    uv-coverage)
    Note: psfmode will be used to clean is imagermode = ''
imagermode -- Advanced imaging e.g. mosaic or Cotton-Schwab clean
    default: imagermode='': Options: '', 'csclean', 'mosaic'
    default '' => psfmode cleaning algorithm used
>>> imagermode='mosaic' expandable parameter(s):
    Image as a mosaic of the different pointings (uses csclean
    style too)
    mosweight -- Individually weight the fields of the mosaic
        default: False; example: mosweight=True
        This can be useful if some of your fields are more
        sensitive than others (i.e. due to time spent
        on-source); this parameter will give more weight to
        higher sensitivity fields in the overlap regions.
    ftmachine -- Gridding method for the image;
        Options: ft (standard interferometric gridding), sd
        (standard single dish) both (ft and sd as appropriate),
        mosaic (gridding use PB as convolution function)
        default: 'mosaic'; example: ftmachine='ft'
    scaletype -- Controls scaling of pixels in the image plane.
        (Not fully implemented...for now only controls
        what is seen if interactive=True...but in the future will
        control the image on which clean components are searched)
        default='SAULT'; example: scaletype='PBCOR'
        Options: 'PBCOR','SAULT'
        'SAULT' when interactive=True shows the residual
        with constant noise across the mosaic. If
        pbcor=False, the final output image is NOT

```

```

        corrected for the PB pattern, and therefore is
        not "flux correct". Division of SAULT
        <im名称>.image by the <im名称>.flux image
        will produce a "flux correct image", can also
        be achieved by setting pbcor=True.
    'PBCOR' uses the SAULT scaling scheme for
        deconvolution, but if interactive=True shows the
        primary beam corrected image; the final PBCOR
        image is "flux correct" if pbcor=True.
>>> imagermode='csclean' expandable parameter(s): Image using the
        Cotton-Schwab algorithm in between major cycles
    cyclefactor -- Change the threshold at which
        the deconvolution cycle will stop, degrid
        and subtract from the visibilities. For
        poor PSFs, reconcile often (cyclefactor=4 or
        5); For good PSFs, use cyclefactor 1.5 to
        2.0. Note: threshold = cyclefactor * max
        sidelobe * max residual.
        default: 1.5; example: cyclefactor=4
    cyclespeedup -- Cycle threshold doubles in this
        number of iterations default: -1;
        example: cyclespeedup=3
        try cyclespeedup = 50 to speed up cleaning
    multiscale -- set of scales to use in deconvolution. If set,
        cleans with several resolutions using hobgom clean. The
        scale sizes are in units of cellsize. So if
        cell='2arcsec', a multiscale scale=10 = 20arcsec. First
        scale should always be 0 (point), we suggest second on
        the order of synthesized beam, third 3-5 times
        synthesized beam, etc. For example if synthesized beam
        is 10" and cell="2", try multiscale = [0,5,15]. Note,
        multiscale is currently a bit slow.
        default: multiscale=[] (standard CLEAN using psfmode algorithm,
        no multi-scale). Example: multiscale = [0,5,15]
>>> multiscale expandable parameter(s): negcomponent -- Stop
        component search when the largest scale has found this
        number of negative components; -1 means continue
        component search even if the largest component is
        negative. default: -1; example: negcomponent=50
    imsize -- Image pixel size (x,y)
        default = [256,256]; example: imsize=[350,350]
        imsize = 500 is equivalent to [500,500]
    cell -- Cell size (x,y)
        default= '1.0arcsec';
        example: cell=['0.5arcsec','0.5arcsec'] or

```

```

    cell=['1arcmin', '1arcmin']
    cell = '1arcsec' is equivalent to ['1arcsec','1arcsec']
    NOTE:cell = 2.0 => ['2arcsec', '2arcsec']
phasecenter -- direction measure or fieldid for the mosaic center
    default: '' => first field selected ; example: phasecenter=6
    or phasecenter='J2000 19h30m00 -40d00m00'
restfreq -- Specify rest frequency to use for output image
    default='' Occasionally it is necessary to set this (for
    example some VLA spectral line data). For example for
    NH_3 (1,1) put restfreq='23.694496GHz'
stokes -- Stokes parameters to image
    default='I'; example: stokes='IQUV';
    Options: 'I','IV','QU','IQUV','RR','LL','XX','YY','RRL','XXYY'
niter -- Maximum number iterations,
    if niter=0, then no CLEANing is done ("invert" only)
    default: 500; example: niter=5000
gain -- Loop gain for CLEANing
    default: 0.1; example: gain=0.5
threshold -- Flux level at which to stop CLEANing
    default: '0.0mJy';
    example: threshold='2.3mJy' (always include units)
           threshold = '0.0023Jy'
           threshold = '0.0023Jy/beam' (okay also)
interactive -- use interactive clean (with GUI viewer)
    default: interactive=False
    example: interactive=True
    interactive clean allows the user to build the cleaning
    mask interactively using the viewer. The viewer will
    appear every npercycle iteration, but modify as needed
    The final interactive mask is saved in the file
    imaname_interactive.mask. The initial masks use the
    union of mask and cleanbox (see below)
>>> interactive=True expandable parameter npercycle -- this is the
    number of iterations between each clean to update mask
    interactively. Set to about niter/5, but can also be
    changed interactively.
mask -- Specification of cleanbox(es), mask image(s), and/or
    region(s) to be used for CLEANing. As long as the image has
    the same shape (size), mask images from a previous
    interactive session can be used for a new execution. NOTE:
    the initial clean mask actually used is the union of what
    is specified in mask and <imaname>.mask default: [] (no
    masking); Possible specification types: (a) Explicit
    cleanbox pixel ranges example: mask=[110,110,150,145] clean
    region with blc=110,100; trc=150,145 (pixel values) (b)

```

```

Filename with cleanbox pixel values with ascii format:
example: mask='mycleanbox.txt' &lt;fieldid blc-x blc-y
trc-x trc-y&gt; on each line
1 45 66 123 124
2 23 100 300 340
(c) Filename for image mask example: mask='myimage.mask'
(d) Filename for region specification (e.g. from viewer)
example: mask='myregion.rgn' (e) Combinations of any of the
above example: mask=[[110,110,150,145], 'mycleanbox.txt',
'myimage.mask', 'myregion.rgn']
uvtaper -- Apply additional uv tapering of the visibilities.
default: uvtaper=False; example: uvtaper=True
&gt;&gt;&gt; uvtaper=True expandable parameters
outertaper -- uv-taper on outer baselines in uv-plane
[bmaj, bmin, bpa] taper Gaussian scale in uv or
angular units. NOTE: uv taper in (klambda) is roughly on-sky
FWHM(arcsec/200)
default: outertaper=[]; no outer taper applied
example: outertaper=['5klambda'] circular taper
FWHM=5 kilo-lambda
outertaper=['5klambda', '3klambda', '45.0deg']
outertaper=['10arcsec'] on-sky FWHM 10"
outertaper=['300.0'] default units are meters
in aperture plane
innertaper -- uv-taper in center of uv-plane
[bmaj,bmin,bpa] Gaussian scale at which taper falls to
zero at uv=0
default: innertaper=[]; no inner taper applied
NOT YET IMPLEMENTED
modelimage -- Name of model image(s) to initialize cleaning. If
multiple images, then these will be added together to
form initial staring model NOTE: these are in addition
to any initial model in the &lt;imagename&gt;.model image file
default: '' (none); example: modelimage='orion.model'
modelimage=['orion.model', 'sdorion.image'] Note: if the
units in the image are Jy/beam as in a single-dish
image, then it will be converted to Jy/pixel as in a
model image, using the restoring beam in the image
header
weighting -- Weighting to apply to visibilities:
default='natural'; example: weighting='uniform';
Options: 'natural', 'uniform', 'briggs',
'superuniform', 'briggsabs', 'radial'
&gt;&gt;&gt; Weighting expandable parameters
For weighting='briggs' and 'briggsabs'

```

```

    robust -- Brigg's robustness parameter
             default=0.0; example: robust=0.5;
             Options: -2.0 to 2.0; -2 (uniform)/+2 (natural)
For weighting='briggsabs'
    noise    -- noise parameter to use for Briggs "abs"
             weighting
             example noise='1.0mJy'
For superuniform/briggs/briggsabs weighting
    npixels -- number of pixels to determine uv-cell size
             for weight calculation
             example npixels=7
restoringbeam -- Output Gaussian restoring beam for CLEAN image
               [bmaj, bmin, bpa] elliptical Gaussian restoring beam
               default units are in arc-seconds for bmaj,bmin, degrees
               for bpa default: restoringbeam=[]; Use PSF calculated
               from dirty beam.
               example: restoringbeam=['10arcsec'] circular Gaussian
                       FWHM 10" example:
                       restoringbeam=['10.0','5.0','45.0deg'] 10"x5"
                       at 45 degrees
pbcor -- Output primary beam-corrected image
         default: pbcor=False; output un-corrected image
         example: pbcor=True; output pb-corrected image (masked outside
         minpb) Note: if you set pbcor=False, you can later
         recover the pbcor image by dividing by the .flux image
         (e.g. using immath)
minpb -- Minimum PB level to use default=0.1; example:
         minpb=0.01 Note: this minpb is always in effect
         (regardless of pbcor=True/False)

```

</example>

</task>

</casaxml>

J.3.2 File task_clean.py

Task clean implementation file.

```

import os
from taskinit import *
from cleanhelper import *

```

```

def clean(vis,imagename,field, spw, selectdata, timerange, uvrange, antenna,
         scan, mode,niter, gain,threshold, psfmode, imagermode, ftmachine,
         mosweight, scaletype, multiscale, negcomponent,interactive,mask,
         nchan,start,width,imsize,cell, phasecenter, restfreq, stokes,
         weighting,robust, uvtaper,outertaper,innertaper, modelimage,
         restoringbeam,pbcor, minpb, noise, npixels, npercycle, cyclefactor,
         cyclespeedup):

#Python script

casalog.origin('clean')

maskimage=''
if((mask==[]) or (mask=='')):
    mask=['']
if (interactive):
    if( (mask=='') or (mask==['']) or (mask==[])):
#try:
if(1):
    imCln=imtool.create()
    imset=cleanhelper(imCln, vis)

    if((len(imagename)==0) or (imagename.isspace())):
        raise Exception, 'Cannot proceed with blank imagename'
    casalog.origin('clean')

    imset.defineimages(imsize=imsize, cell=cell, stokes=stokes,
                      mode=mode, spw=spw, nchan=nchan,
                      start=start, width=width,
                      restfreq=restfreq, field=field,
                      phasecenter=phasecenter)

    imset.datselweightfilter(field=field, spw=spw,
                             timerange=timerange,
                             uvrange=uvrange,
                             antenna=antenna, scan=scan,
                             wgttype=weighting,
                             robust=robust,
                             noise=noise, npixels=npixels,
                             mosweight=mosweight,
                             innertaper=innertaper,
                             outertaper=outertaper)

```

```

if(maskimage==''):
    maskimage=imagename+'.mask'
imset.makemaskimage(outputmask=maskimage,imagename=imagename,
                    maskobject=mask)

###define clean alg
alg=psfmode
if(multiscale==[0]):
    multiscale=[]
if((type(multiscale)==list) and (len(multiscale)>0)):
    alg='multiscale'
    imCln.setscales(scalemethod='uservector',
                   uservector=multiscale)
if(imagermode=='csclean'):
    alg='mf'+alg
if(imagermode=='mosaic'):
    if(alg.count('mf') <1):
        alg='mf'+alg
    imCln.setoptions(ftmachine=ftmachine, padding=1.0)
    imCln.setvp(dovp=True)
###PBCOR or not
sclt='SAULT'
if((scaletype=='PBCOR') or (scaletype=='pbcor')):
    sclt='NONE'
    imCln.setvp(dovp=True)
else:
    if(imagermode != 'mosaic'):
        ##make a pb for flux scale
        imCln.setvp(dovp=True)
        imCln.makeimage(type='pb', image=imagename+'.flux')
        imCln.setvp(dovp=False)

##restoring
imset.setrestoringbeam(restoringbeam)
###model image
imset.convertmodelimage(modelimages=modelimage,
                        outputmodel=imagename+'.model')

####after all the mask shenanigans...make sure to use the
####last mask
maskimage=imset.outputmask
if((imagermode=='mosaic')):
    imCln.setmfcontrol(stoplargenegatives=negcomponent,scaletype=sclt,

```

```

        minpb=minpb,cyclefactor=cyclefactor,
        cyclespeedup=cyclespeedup,
        fluxscale=[imagename+'.flux'])
else:
    imCln.setmfcontrol(stoplargenegatives=negcomponent,
        cyclefactor=cyclefactor, cyclespeedup=cyclespeedup)

imCln.clean(algorithm=alg,niter=niter,gain=gain,
    threshold=qa.quantity(threshold,'mJy'),
    model=[imagename+'.model'],
    residual=[imagename+'.residual'],
    image=[imagename+'.image'],
    psfimage=[imagename+'.psf'],
    mask=maskimage, interactive=interactive,
    npercycle=npercycle)
imCln.close()
presdir=os.path.realpath('.')
newimage=imagename
if(imagename.count('/') > 0):
    newimage=os.path.basename(imagename)
    os.chdir(os.path.dirname(imagename))
result          = '\ ' + newimage + '.image' + '\ '
fluxscale_image = '\ ' + newimage + '.flux' + '\ '
if (pbcor):
    if(sclt != 'NONE'):
        ##otherwise its already divided
        ia.open(newimage+'.image')

        pixmask = fluxscale_image+'>'+str(minpb);
        ia.calcmask(pixmask,asdefault=True);

        pixels='iif('+ fluxscale_image+'>'+str(minpb)+','
            + result+'/' +fluxscale_image+', 0)'
        ia.calc(pixels=pixels)
        ia.close()
else:
    ## people has imaged the fluxed corrected image
    ## but want the
    ## final image to be non-fluxed corrected
    if(sclt=='NONE'):
        ia.open(newimage+'.image')
        result=newimage+'.image'
        fluxscale_image=newimage+'.flux'
        pixels=result+'*'+fluxscale_image
        ia.calc(pixels=pixels)

```



```
                ia.close()
            os.chdir(presdir)

        del imCln

#     except Exception, instance:
#         print '*** Error *** ',instance
#         raise Exception, instance
```