Visibility Data Inspection

Loránt Sjouwerman, NRAO

Goal: prepare for (easier) calibration

Examine the visibility data, start from archive download

- Science Data Model (SDM) and Binary Data Format (BDF)
  - SDM = metadata; sources, antennas, correlator, weather, etc
  - BDF = single scan data ("missing BDF": scan lost)
- Convert to data reduction package format, here **CASA**
- Retrace observing schedule, observing conditions
  - Data structure, scan intents, antennas, operator logs (cf. OPT)
- Remove (flag) problematic data, items for watchlist
  - Known issues (off-source, unused antenna)
  - Unforeseen issues (RFI, unnoticed by operator)
  - Questionable data (may need flagging later, or not..)

CASA

- See previous lecture…
- Documentation is available at
  - http://casa.nrao.edu/
- Training material is available at
  - http://casaguides.nrao.edu/
- For help use the NRAO helpdesk at
  - http://help.nrao.edu/
- CASA 6.2.1 will be used at this workshop
    - Type "casa –ls" to find the exact version string
    - Then "casa [-r <version-string>]" to start

Loading the data: *importasdm*

Goal: convert SDM/BDF into CASA visibility data format

- Skip if CASA MS ("measurement set") format is downloaded
- Task *import**asdm*** converts SDM/BDF to MS for CASA processing
  - *import**vla*** for older data; more options like *import**fits*** ..
- *importasdm* understands and applies VLA online flags:

	default importasdm		*set task and default parameters*
	inp		*review parameter options*
Read:	asdm	=	'SDM/BDF_directory_string'
Write:	vis	=	'CASA_MS_directory_string'
	ocorr_mode	=	'co'
	applyflags	=	True
	inp		*review parameter selections*
	go		*execute task*

Loading the data: *flagdata*

If applyflags=False (default) a FLAG_CMD table is created

- Examine, list/plot (before applying by default) using *flagcmd*
  - Also used for other flagging operations, see later in this talk
- Additional useful a-priori flagging uses *flagdata* (also see later)

	default flagdata		*set task and default parameters*
Read/write:	vis	=	'CASA_MS_directory_string'
Shadowing:	mode	=	'shadow'
	inp	=	*review parameter selections*
	go	=	*execute task*
Pure zero data:	mode	=	'clip'
	correlation		'ABS_ALL'
	clipzeros		True
	go		*execute task*

Examine your observation/data

Goal: know your data content and find/fix possible issues (that will give problems) before starting the calibration

- Operator observing log (email to observers)
  - http://www.vla.nrao.edu/cgi-bin/oplogs.cgi
- Plotting the antenna positions: *plotants*
  - Array configuration, possible reference antennas
- Observing summary: *listobs*
  - Calibrator sources, scan intents, spectral configurations, etc..
- Plotting/displaying/editing data: *plotms*
  - Examine, assess, act: *(visibility) to be or not to be (retained)*
- (pipeline operations weblog might be useful too if available)

Operator observing log

Goal: first impression about the observation

- Array status, weather, antenna problems etc, downtime measure

VLA OBSERV[illegible] **2018-10-04_0541_TDRW0001**

Obser[illegible] Date:	04-Oct-2018	**Project:**	TDRW0001	**# Subarrays:** 1	**Observation Type:** Test
Con[illegible]uration:	D	**Observer(PI):**	Dr Emmanuel Mo[illegible]an		**Band(s) Used:** C S
De[illegible]mmissioned:	27	**SBID(s):**	35624494		
		Source File(s):	TDRW0001_sb35[illegible]494_1_1		
		Observer E-mail:	emomjian@n[illegible].edu		
		Operator(s):	Kenn[illegible]son		

Adobe PDF version of this log is located at: http://www.vla.nrao.edu/operators/logs/
Visibility data is updated each day at IAT/UT midnight and is available from the online archive at: https://archive.nrao.edu

Time (U[illegible]	Dew Point (C)	Temp. (C)	Wind Speed & Direction (avg)	Bar. Pressure (mbars)	[illegible] Phase (degs)	Remarks
04[illegible]:44:10	6.2	17.2	SW at 6.7 m/s	789.7	4.3	Sky cover 2[illegible] Cumuliform clouds.
[illegible]ct 6:46:31	7.0	14.3	SW at 4.7 m/s	789.4	4.7	Sky cover 10%. [illegible] Stratiform clouds.

Number of antennas used: 27

Start Time	End Time	Comments/Outages	Form #	#Ants	Down Time (in minutes)
04Oct 5:41:30		Starting project TDRW0001.			
04Oct 5:41:30		The band(s) used is(are): C S.			
04Oct 5:44:06		On source 0137+331=3C48 [illegible] available antennas.			
04Oct 5:41:30		To access your data f[illegible] NRAO archive visit:			
		https://scienc[illegible].edu/facilities/vla/archive.			
		All VLA [illegible]e data are processed through the VLA calibration pipeline. Details			
		[illegible] https://science.nrao.edu/facilities/vla/data-processing/pipeline.			
		[illegible] or further questions please use the NRAO helpdesk at:			
		https://science.nrao.edu/observing/helpdesk.			
04Oct 5:41:30		Note: To support our ongoing RFI monitoring efforts, any feedback from your			
		program on RFI can be sent to: nrao-rfi@nrao.edu.			
		The key information to provide is:			
		- Observation/project code			
		- Frequency and Time of the observations			
		- The characteristics of the RFI signal, in particular if it is continuous or			

VLA OBSERVING LOG **2018-10-04_0541_TDRW0001**

		intermittent?			
		- If possible, a spectrum of the RFI should be included in the e-mail.			
		Thanks very much for your support; this [illegible]			
		updated on the [illegible]			
		[illegible].nrao.edu/facilities/vla/docs/manuals/obsguide/modes/rfi/			
04Oct 5:41:30		Antennas in the D-array may be shadowed at low elevations. If shadowing			
		occurs, sensitivity will be affected.			
04Oct 5:41:30		NOTE!: The VLA is still recovering from a long power outage, and these data may			
		have unusual artifacts, missing antennas or IFs, ect., in them. NRAO staff will			
		examine the data closely after observing to determine if they meet the criteria for			
		a successful observation.			
04Oct 5:41:30	04Oct 8:32:51	Antenna(s) 5 (Data: Corrupted): FRONT END [illegible]959		0.97	165.4
		S-band receiver cooling after work performed, currently 65/177K			
04Oct 5:41:30	04Oct 8:32:51	Antenna(s) 12 (Data: Corrupted): CR[illegible]ENICS	PM	0.03	5.8
		C-band receiver warm for cold head replacement			
04Oct 5:57:00		Your new operator(s) is(are): Sam Gilmore			
Project End Time			**[illegible]otal Project Time (minutes x 27 ants.)**	**Down Time % of Total Time**	**Total Down Time**
04Oct 8:32:51	**End of project TDRW0001**		**4626.5**	**3.7%**	**171.2**

Choosing the reference antenna(s)

Goal: good 'refant' reduces loss due to flagging/calibration

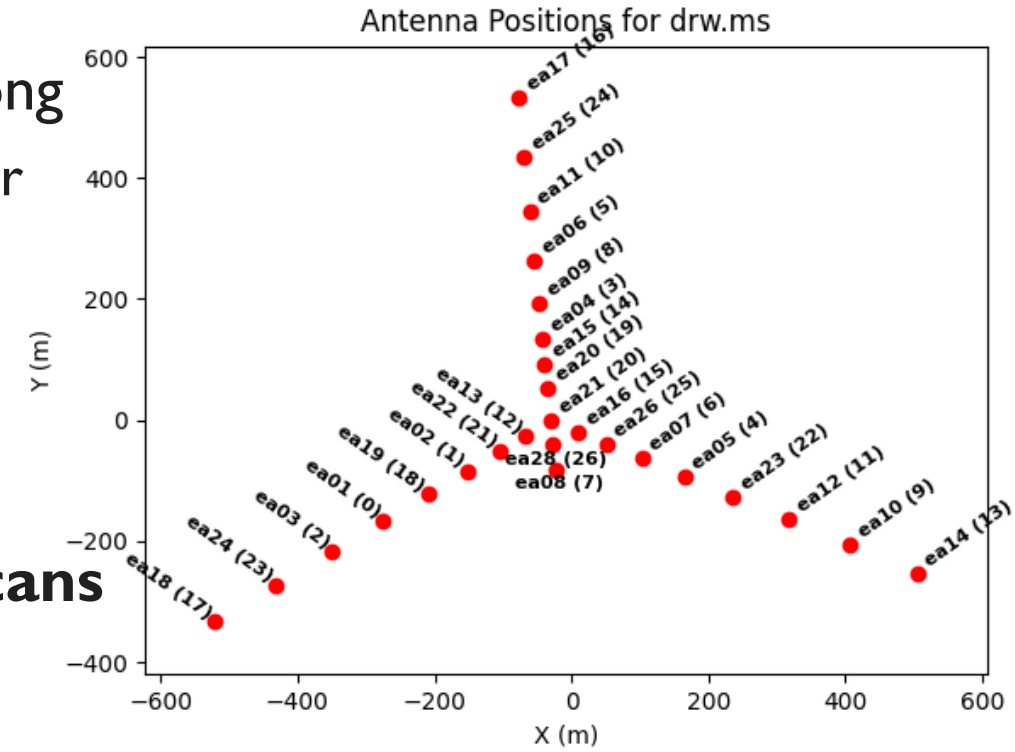
- Use *plotants*:
- Baselines/uv-dist not too long
  - **Near** physical array center
    - Use *listobs/vishead*
  - e.g. 08-pad: [WNE]08
    - **A**: inner antennas
    - **D**: no shadowing !
- Good on **all calibrator scans**
    - Target scans irrelevant
    - No antenna "issues"
    - Clean from RFI

default plotants		*set task and default parameters*
inp		*review parameter options*
vis	=	'CASA_MS_directory_string'
antindex	=	True
inp		*review parameter selections*
go		*execute task*

Choosing the reference antenna(s)

Goal: good 'refant' reduces loss due to flagging/calibration

- Use *plotants*:
- Baselines/uv-dist not too long
  - **Near** physical array center
    - Use *listobs*/*vishead*
  - e.g. 08-pad: [WNE]08
    - **A**: inner antennas
    - **D**: no shadowing !
- Good on **all calibrator scans**
    - Target scans irrelevant
    - No antenna "issues"
    - Clean from RFI



Observing structure

Goal: identify key strategy, scan outline, calibration hints

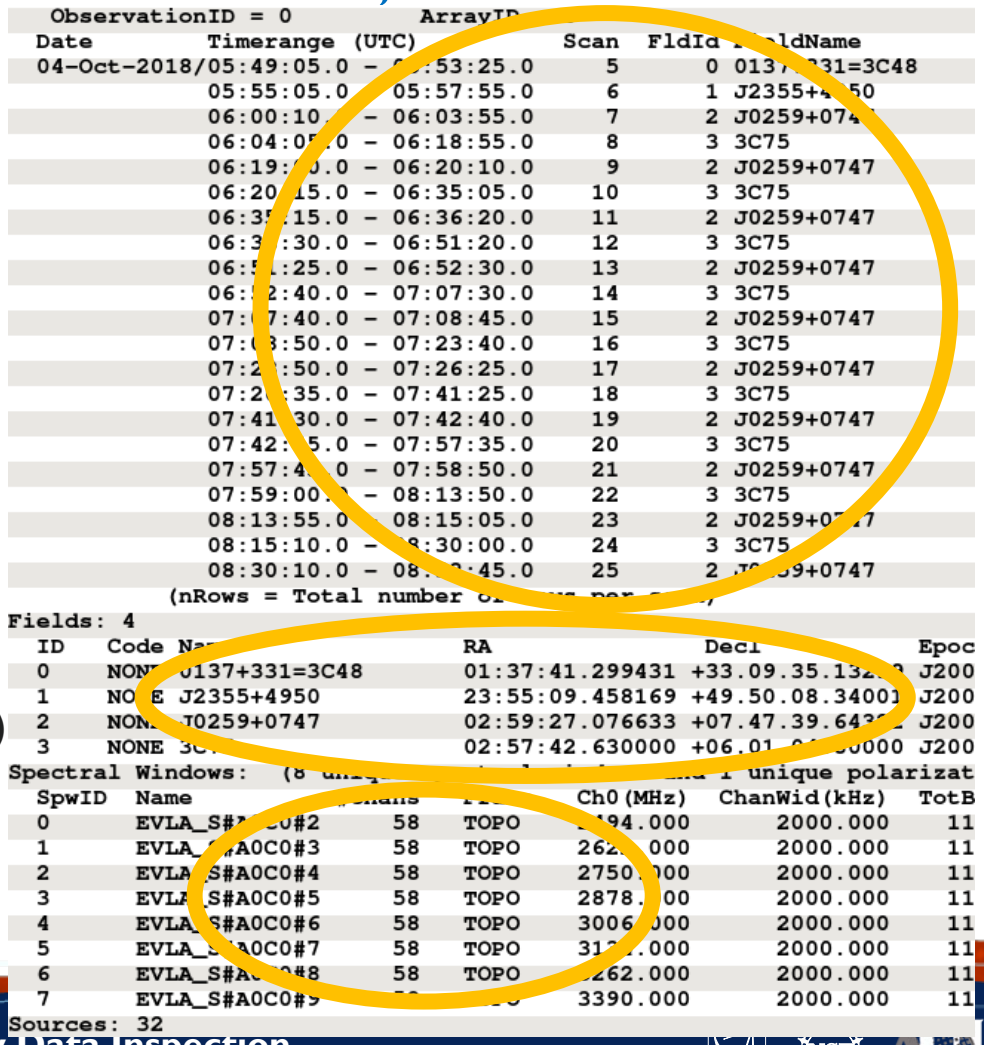
- Use *listobs* (not *vishead*):
- Source (i.e. field) names
- Scan intents: calibrators
  - Flux density
  - Bandpass
  - Gain (phase and amp!)
- Correlator setup
  - Spectral windows
    - (IFs/subbands, channels)
  - Polarization
  - Doppler setting

default listobs		*set task and default parameters*
vis	=	'CASA_MS_directory_string'
inp		*review parameter selections*
go		*execute task*
NOTE:		
listfile	=	'output_filename_string'

Observing structure

Goal: identify key strategy, scan outline, calibration hints

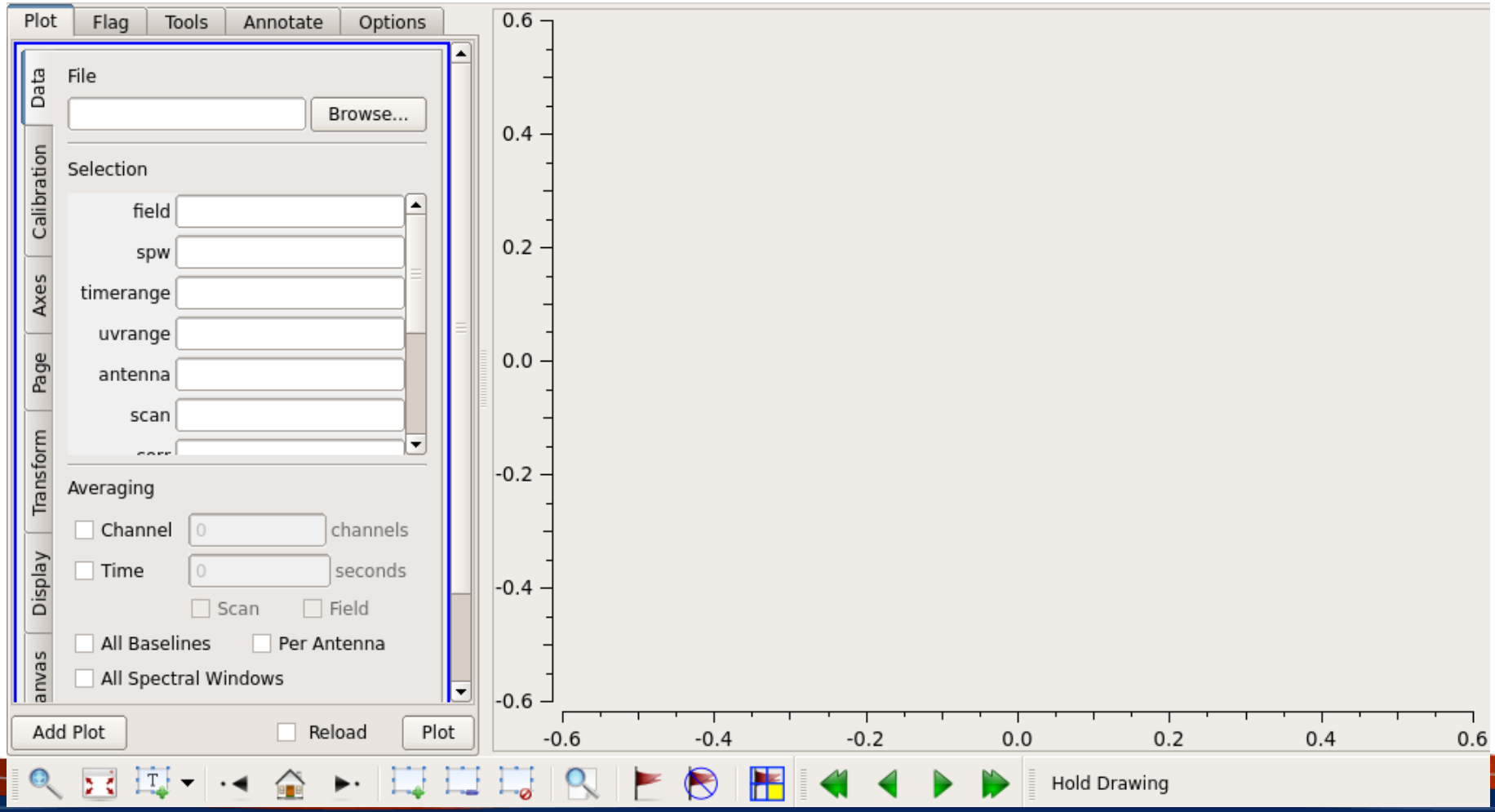
- Use *listobs* (not *vishead*):
- Source (i.e. field) names
- Scan intents: calibrators
  - Flux density
  - Bandpass
  - Gain (phase and amp!)
- Correlator setup
  - Spectral windows
    - (IFs/subbands, channels)
  - Polarization
  - Doppler setting



```
  ObservationID = 0         ArrayID
  Date          Timerange (UTC)          Scan  FldId FieldName
  04-Oct-2018/05:49:05.0 - 05:53:25.0       5      0 0137+331=3C48
              05:55:05.0 - 05:57:55.0       6      1 J2355+4950
              06:00:10.0 - 06:03:55.0       7      2 J0259+0747
              06:04:05.0 - 06:18:55.0       8      3 3C75
              06:19:00.0 - 06:20:10.0       9      2 J0259+0747
              06:20:15.0 - 06:35:05.0      10      3 3C75
              06:35:15.0 - 06:36:20.0      11      2 J0259+0747
              06:36:30.0 - 06:51:20.0      12      3 3C75
              06:51:25.0 - 06:52:30.0      13      2 J0259+0747
              06:52:40.0 - 07:07:30.0      14      3 3C75
              07:07:40.0 - 07:08:45.0      15      2 J0259+0747
              07:08:50.0 - 07:23:40.0      16      3 3C75
              07:23:50.0 - 07:26:25.0      17      2 J0259+0747
              07:26:35.0 - 07:41:25.0      18      3 3C75
              07:41:30.0 - 07:42:40.0      19      2 J0259+0747
              07:42:45.0 - 07:57:35.0      20      3 3C75
              07:57:40.0 - 07:58:50.0      21      2 J0259+0747
              07:59:00.0 - 08:13:50.0      22      3 3C75
              08:13:55.0 - 08:15:05.0      23      2 J0259+0747
              08:15:10.0 - 08:30:00.0      24      3 3C75
              08:30:10.0 - 08:31:45.0      25      2 J0259+0747
          (nRows = Total number of rows per scan)
Fields: 4
  ID   Code Name                RA               Decl           Epoc
  0    NONE 0137+331=3C48       01:37:41.299431 +33.09.35.132  J200
  1    NONE J2355+4950          23:55:09.458169 +49.50.08.3400 J200
  2    NONE J0259+0747          02:59:27.076633 +07.47.39.643  J200
  3    NONE 3C75                02:57:42.630000 +06.01.00.000  J200
Spectral Windows:  (8 unique frequency setups) and 1 unique polarizat
  SpwID  Name                #Chans   Frame  Ch0(MHz)   ChanWid(kHz)  TotB
  0      EVLA_S#A0C0#2        58       TOPO    2194.000      2000.000     11
  1      EVLA_S#A0C0#3        58       TOPO    2262.000      2000.000     11
  2      EVLA_S#A0C0#4        58       TOPO    2750.000      2000.000     11
  3      EVLA_S#A0C0#5        58       TOPO    2878.000      2000.000     11
  4      EVLA_S#A0C0#6        58       TOPO    3006.000      2000.000     11
  5      EVLA_S#A0C0#7        58       TOPO    3134.000      2000.000     11
  6      EVLA_S#A0C0#8        58       TOPO    3262.000      2000.000     11
  7      EVLA_S#A0C0#9        58       TOPO    3390.000      2000.000     11
Sources: 32
```

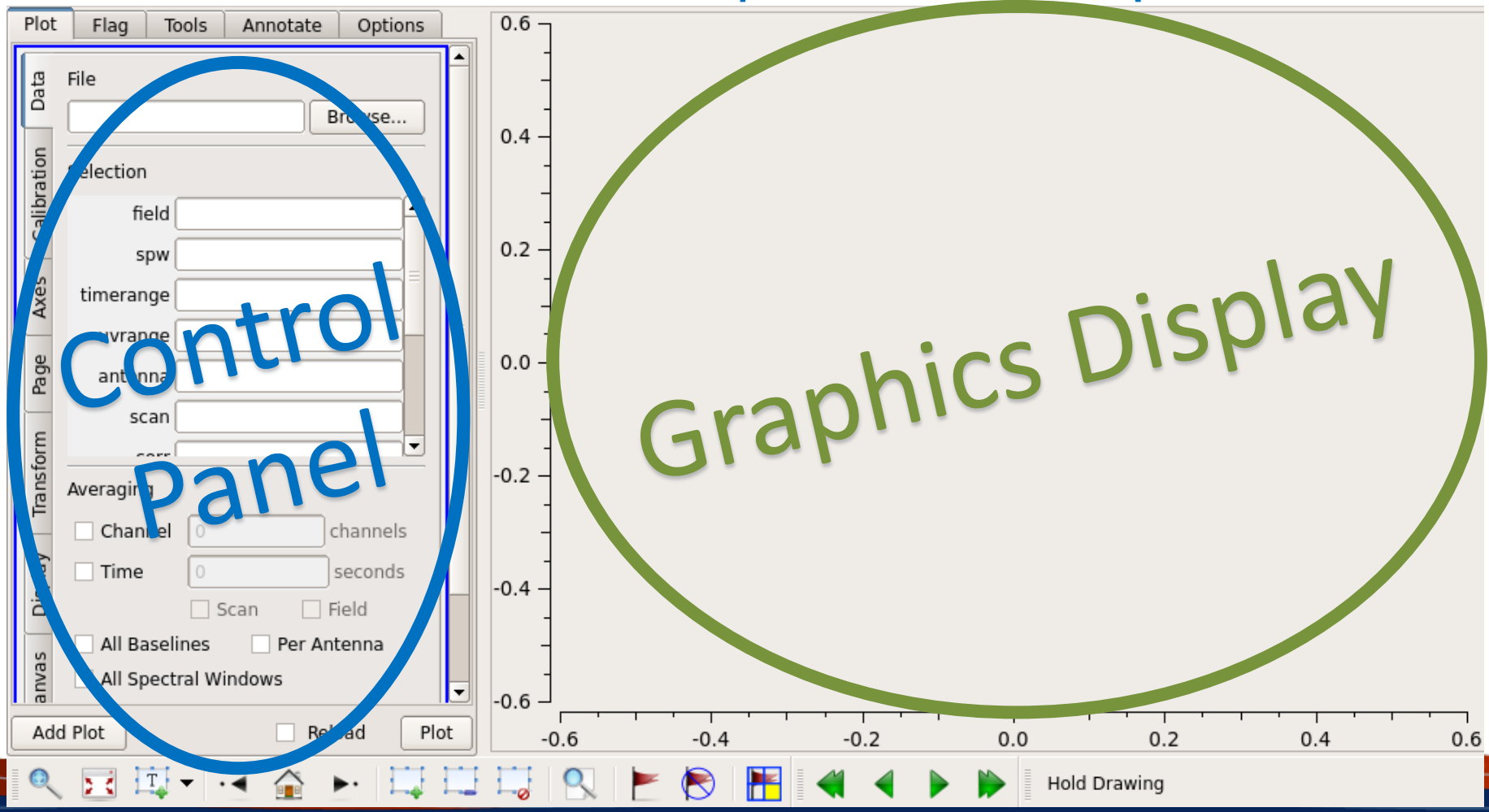
Examine MS visibility data: *plotms*

Intermezzo: introduction to *plotms* GUI and panels:



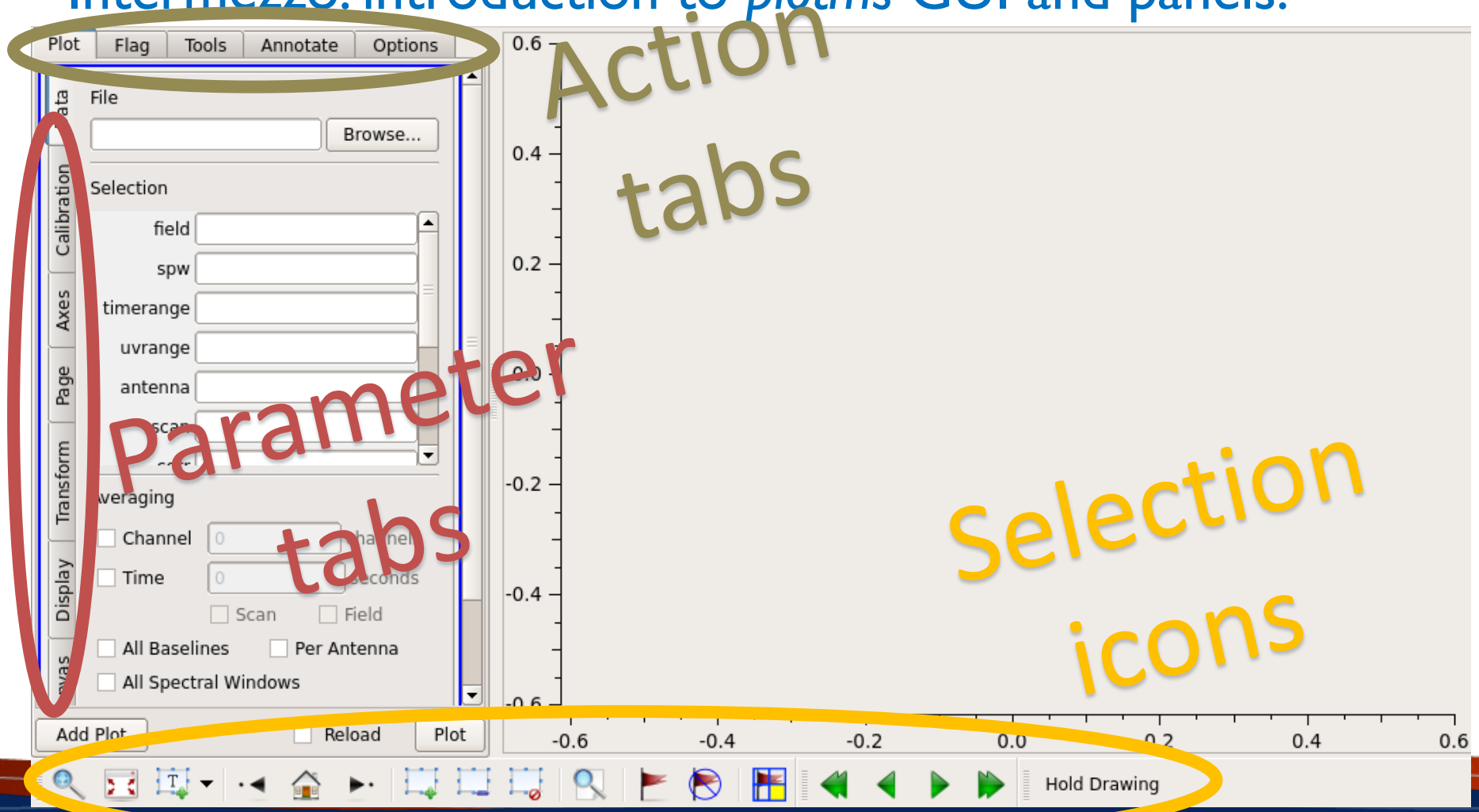
Examine MS visibility data: *plotms*

Intermezzo: introduction to *plotms* GUI and panels:



Examine MS visibility data: *plotms*

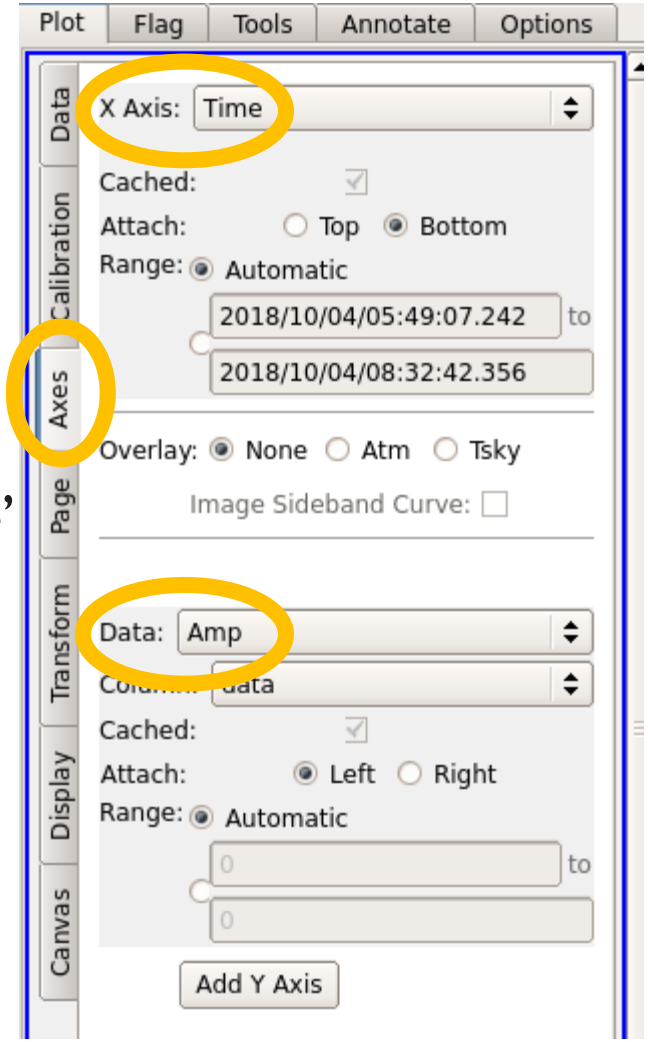
Intermezzo: introduction to *plotms* GUI and panels:



Control panel parameters

Axes options

- Typical **X Axis** selections (not all):
  - 'time'/'scan'/'field'
  - 'frequency'(freq)/'spw'
  - 'channel'(chan)/'velocity'(vel)
  - 'antenna1'(ant1)/'antenna2'(ant2)/'baseline'
  - 'UVdist'/'UVwave'
- Typical **Data** (i.e. Y-axis) selections:
  - 'amplitude'(amp)
  - 'phase'
  - 'real'/'imaginary'(imag)
  - 'weight'(wt)/'flag'



Control panel parameters

Display options

- Typical **Colorize** selections:

  For time-sequence-like x-axis

  – Field, scan

  – Antenna1/2, baseline

  For frequency-order-like x-axis

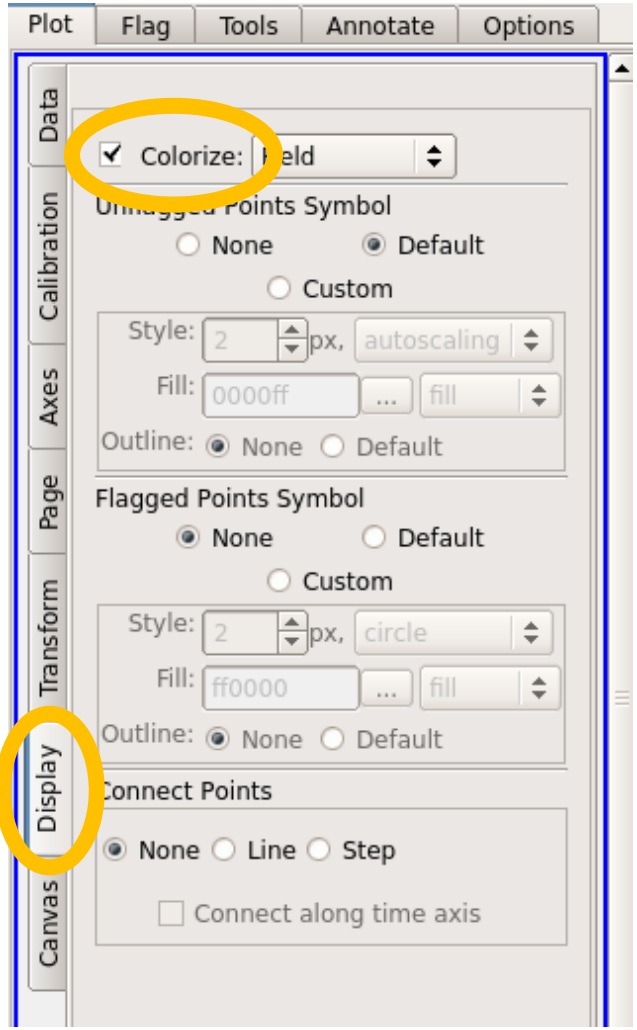
  – spw

  – Correlation (i.e. polarization)

  – Antenna1/2, baseline

  For array-configuration-like x-axis

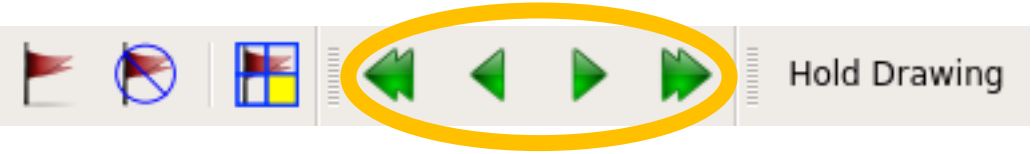
  – Antenna1/2, baseline, spw, field, correlation



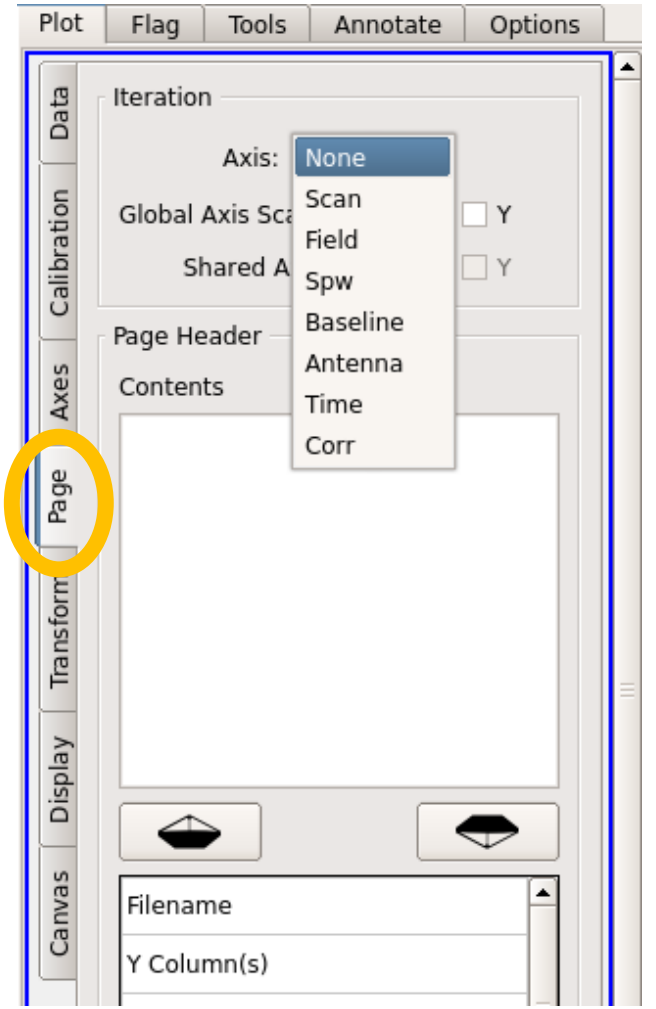
Control panel parameters

Page options

- In multi-page mode, when plots are distributed over more than one plot, one can cycle through the plots using the page-selection tool icons:



- Typical **multi-page** selections:
  - Antenna
  - Correlation
  - Field
  - spw



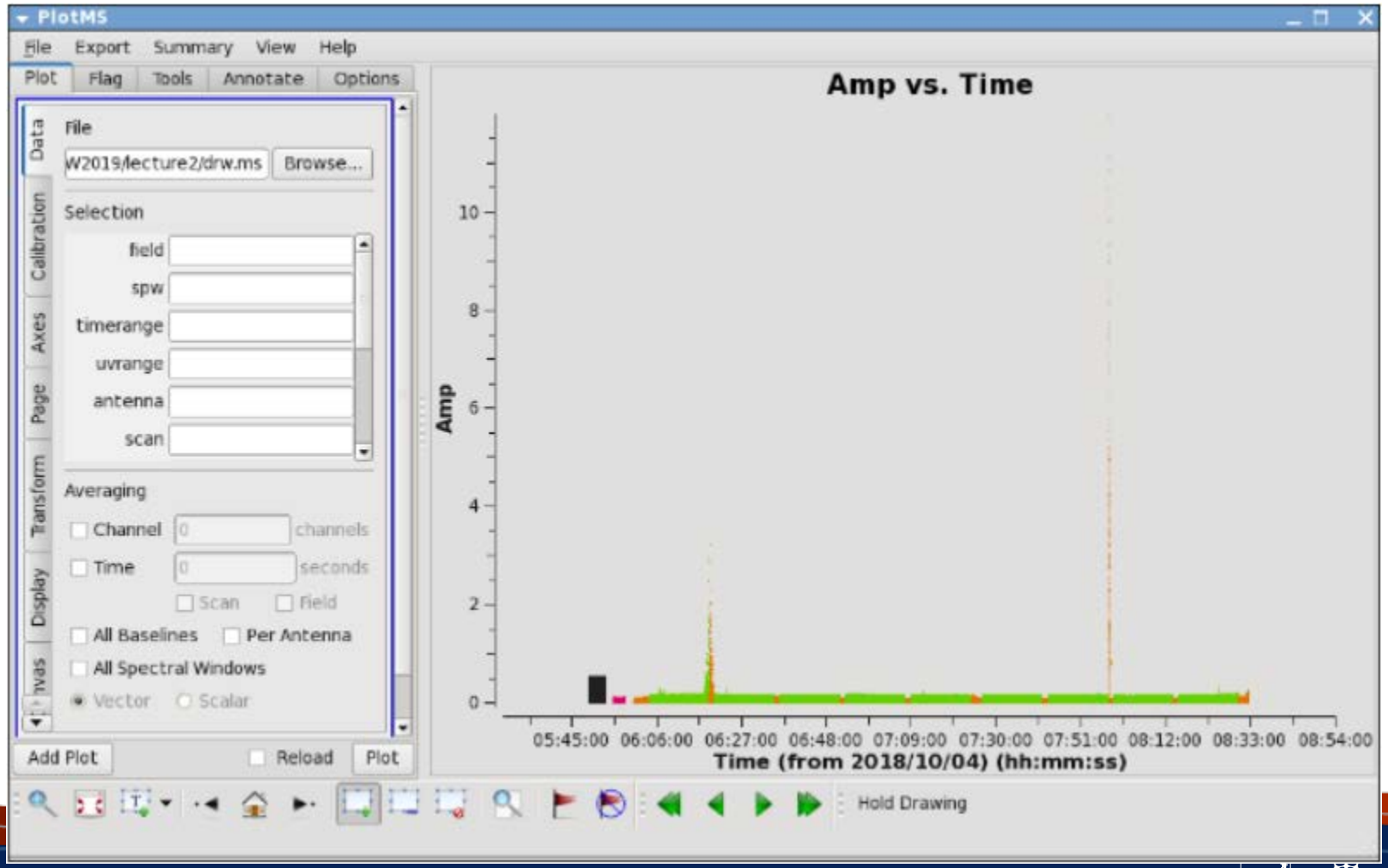
Examine, assess, act

Goal: remove potentially problematic data, identify other potential problems that may influence calibration/imaging

- Overall sense of calibrators/intents and target visibilities, frequency setup, observing conditions, instrumental response
- Where to expect potentially bad data, note the reference antenna!
  - Ill-performing antennas/baselines/samplers/polarizations (instrument)
  - In time domain:
    - Start of scans (not yet on-source, settling time)
    - Bad weather/pointing/RFI-bursts (observing conditions)
  - In frequency domain:
    - Bandpass, subband edges and baseband roll-off (channel sensitivity)
    - RFI-frequencies – not your line!

Time/scan/field dependent issues

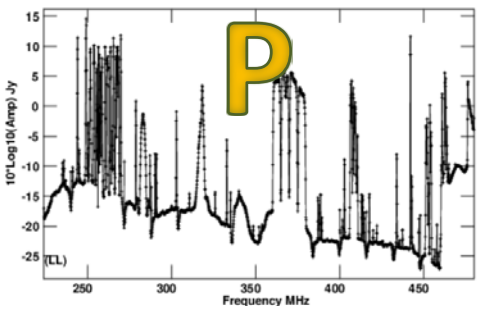
Example: xaxis='time', yaxis='amp', coloraxis='field'

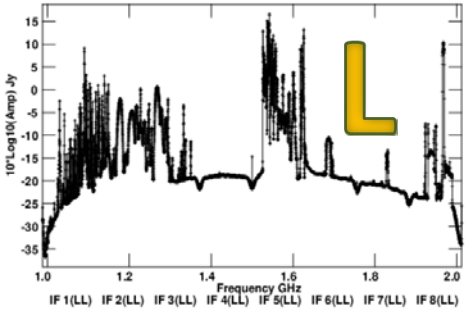


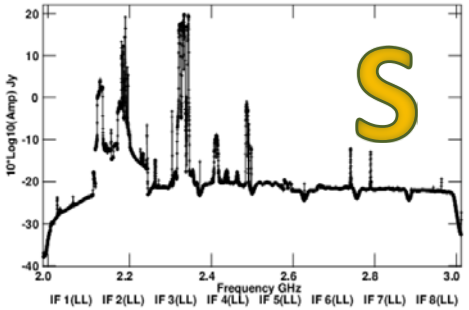
Radio frequency interference (RFI)

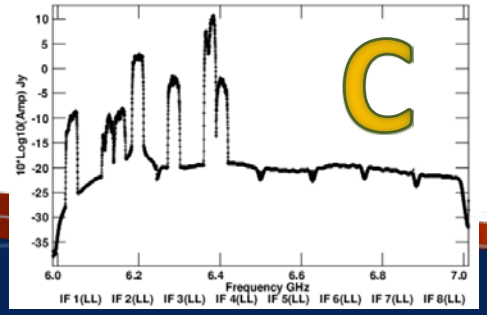
Intermezzo: RFI awareness

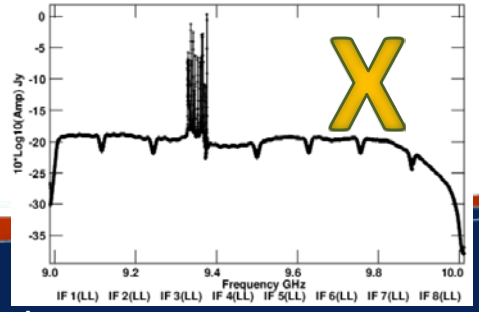
- Observations, in particular in the observing bands 4, P, L, S, C, but also X and Ku, **will** be affected by time/frequency dependent RFI
- VLA RFI information is available at http://go.nrao.edu/vla-rfi
  - Contains RFI listings per frequency band and array configuration
  - Shows (snapshot) spectra of various RFI sweeps between 1-50 GHz

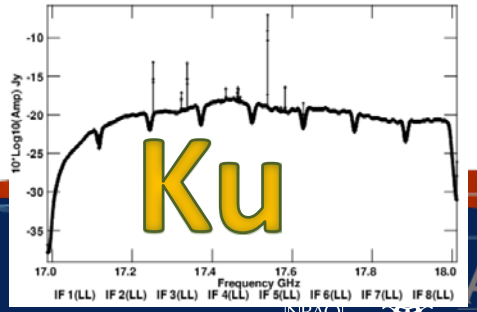






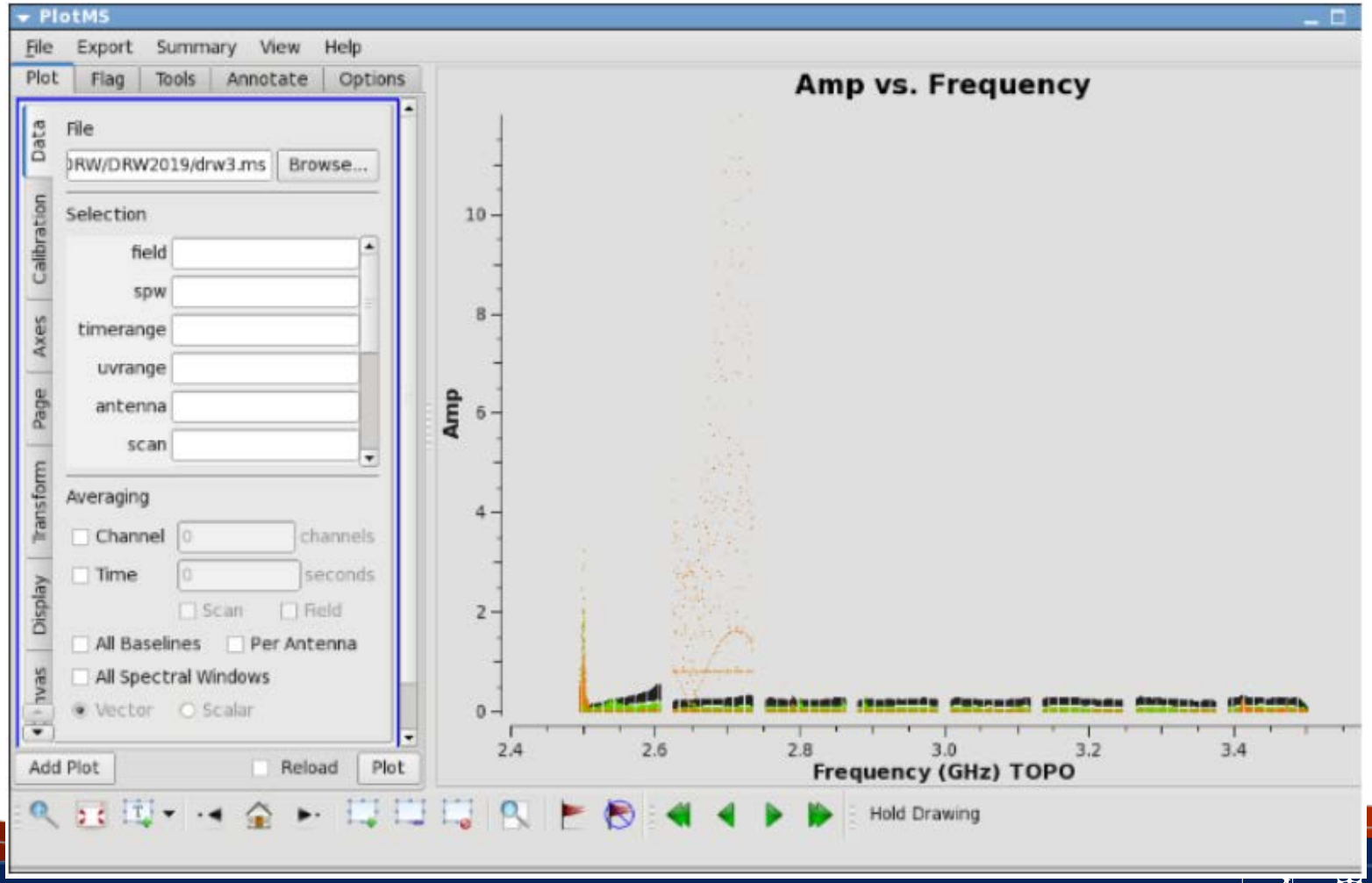






Frequency/spw dependent issues

Example: xaxis='frequency', yaxis='amp', coloraxis='field'



Editing (flagging or unflagging) data

Goal: introduce different flagging tasks and methods

- *flagdata*: All purpose flagging based on selection and algorithms
  - In particular automated RFI flagging operations RFLAG and TFCROP
- *flagcmd*: All purpose flagging based on commands
  - See earlier shadowing and pure zero visibility examples
- *plotms*: interactive flagging using graphical user interface

- Note that the VLA operator log has useful information but may not capture all (or too many) possible items to flag, e.g.
  - Antennas not participating in array may need manual flagging
  - "low amplitude"/"weak": pointing/baseband error or recoverable?

Editing (flagging or unflagging) data

Goal: understand flagging and flag-backups in CASA

- Each visibility is either **used** (not flagged) or **flagged** (ignored)
  - A flag-column contains a False (default) or True bit for each visibility
  - "Applying flagging" converts this bit to True, **regardless** the state it was in before; there is no record keeping of the original value
  - "Unflagging" **resets** this bit to False, with no knowledge of the original value – if it was previously flagged this will be undone too!
- **Most** flagging tasks – **but not *plotms*! -** allow to save a copy of the previous state, before applying new flags, in a **flag-backup** file
- *flagmanager* restores/creates a/the previous/current flagging state
  - When planning to flag with *plotms*, maybe make a flag-backup first?
  - Sequential flag-backup files restore to sequential stages in process

Using *flagdata*

Different modes and options to use in *flagdata*

- *manual* – flag based on specific selection of parameters (default)
- *list* – apply a list of flagging commands in a file or list of files/string
- *clip/quack* – clip on amplitude or (scan)time range values
- *shadow/elevation* – flag on antenna geometry
- *tfcrop* – algorithmic removal of outliers in time-frequency plane
- *rflag* – algorithmic removal of outliers with respect to local RMS
- *extend/antint* – expand flags along different axis (time, freq, ants, ..)
- summary – list what is flagged
- unflag – reset flag-column bits to False (but be aware and careful!)

Using *flagcmd*

Different modes and options to use in *flagcmd*

- Table – flag items in the FLAG_CMD table (default)
- List - apply a list of flagging commands in a file
- Xml – apply on-line flags as listed in the Flag.xml file of the MS

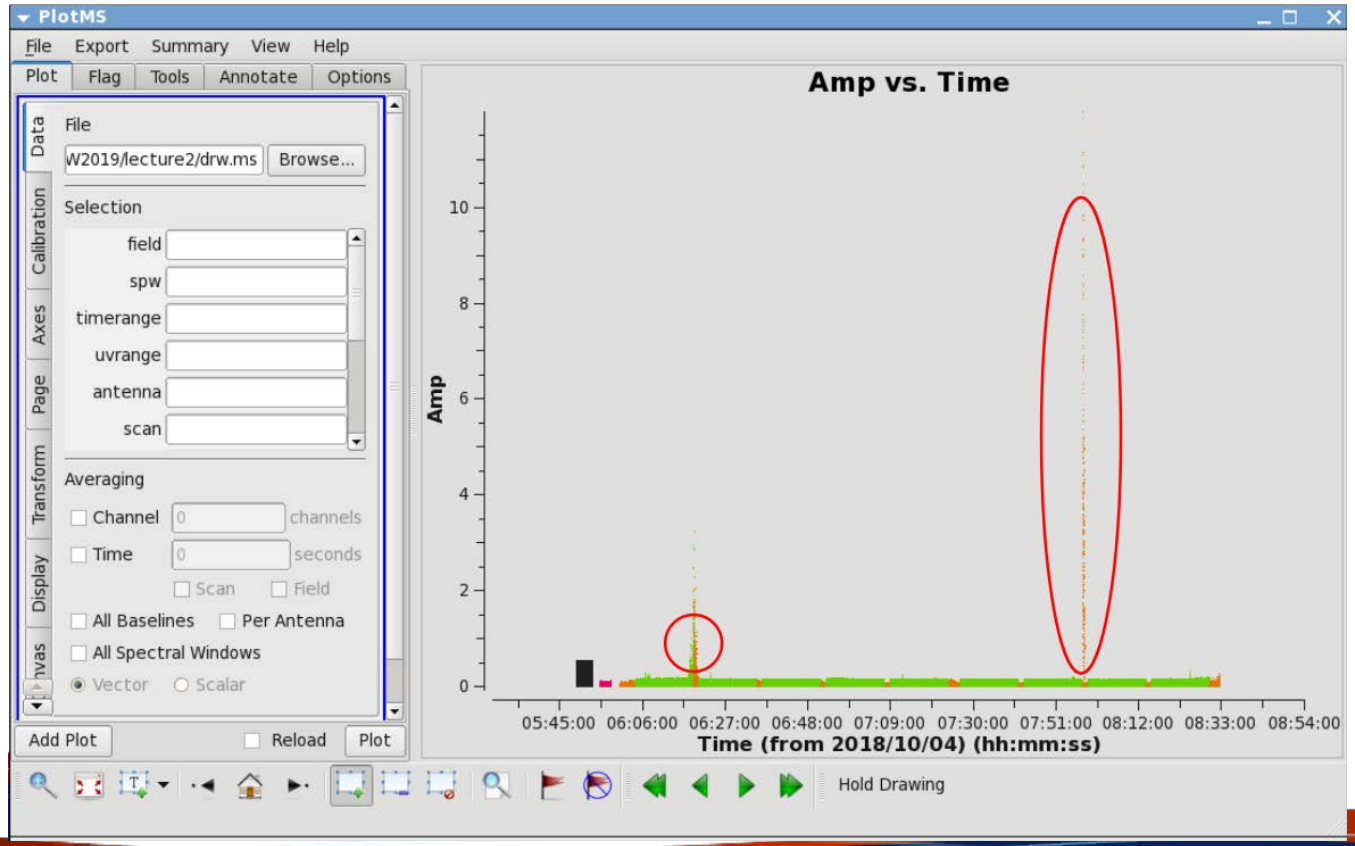
- *flagcmd* can make a plot of flags but lacks runtime displays

- Both *flagcmd* and *flagdata* allow saving to FLAG_CMD file
- Both *flagcmd* and *flagdata* have unflag/unapply operations
- Both *flagcmd* and *flagdata* also operate on **calibration tables**

Example *plotms* editing

Goal: **look for outliers**, here amplitude versus time

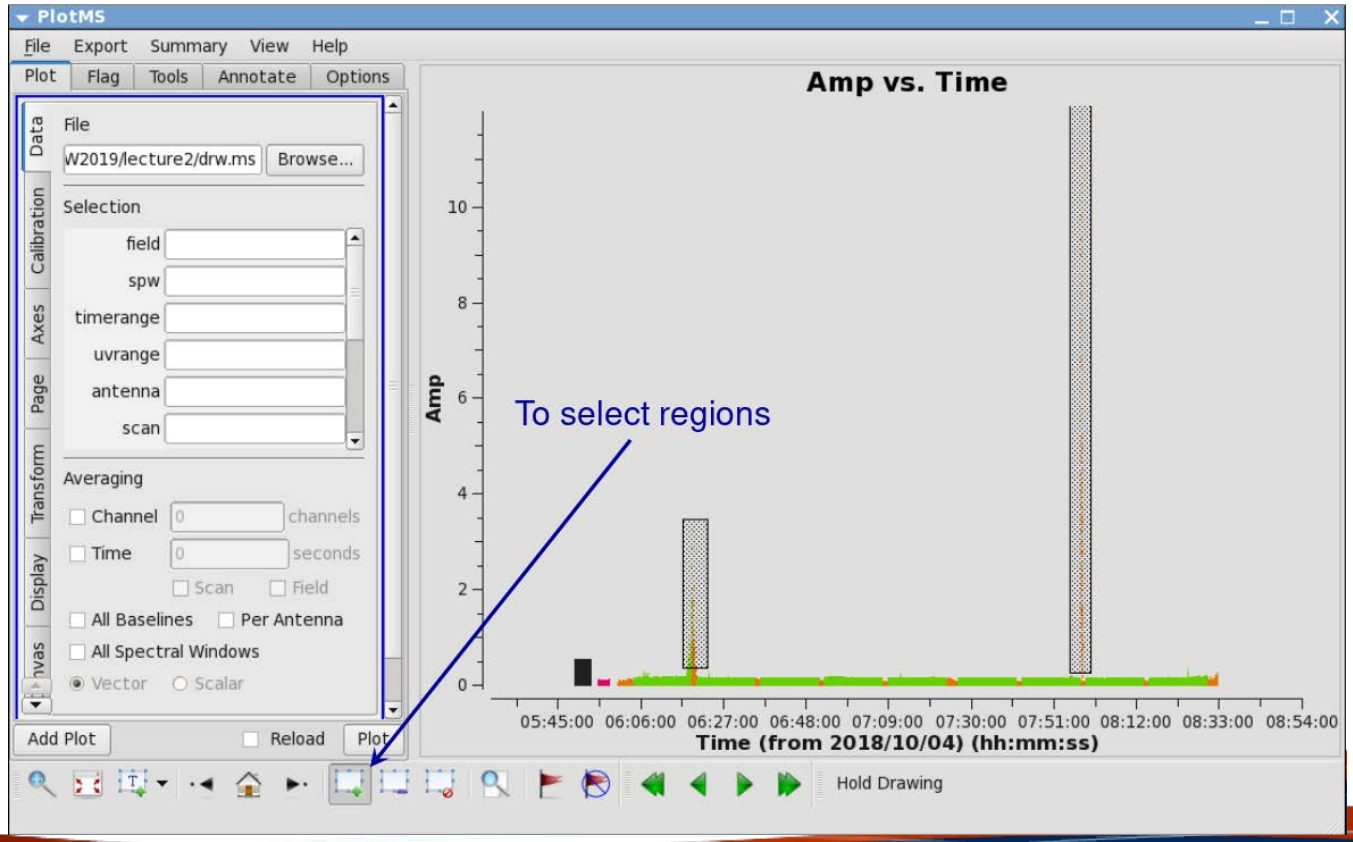
- Two obvious times where probably "something is going on"



Example *plotms* editing

Goal: **select outliers with** select **tool icon**

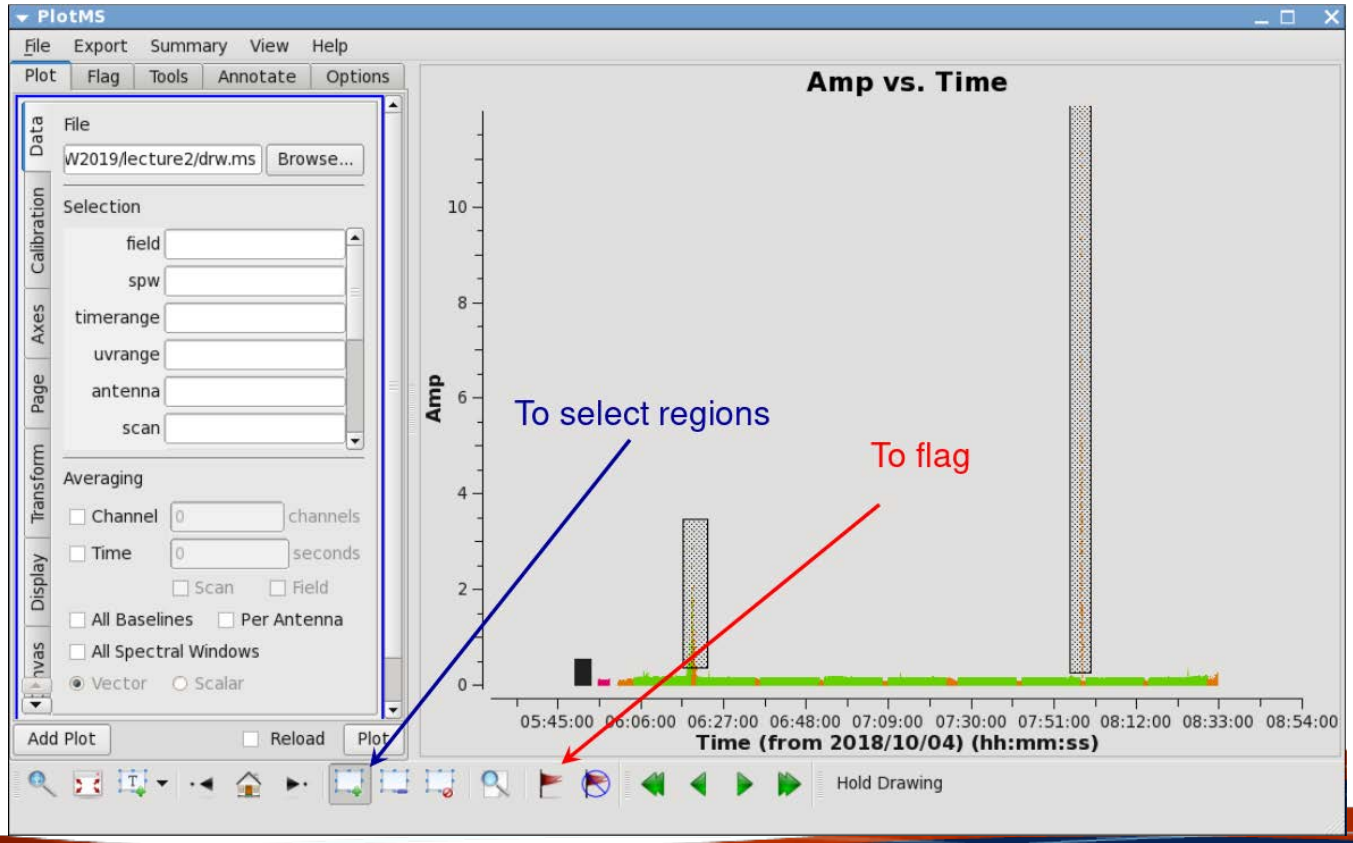
- Use mouse to draw; selections show up as shaded regions



Example *plotms* editing

Goal: **remove outliers with** flag **tool icon** (but wait..!)

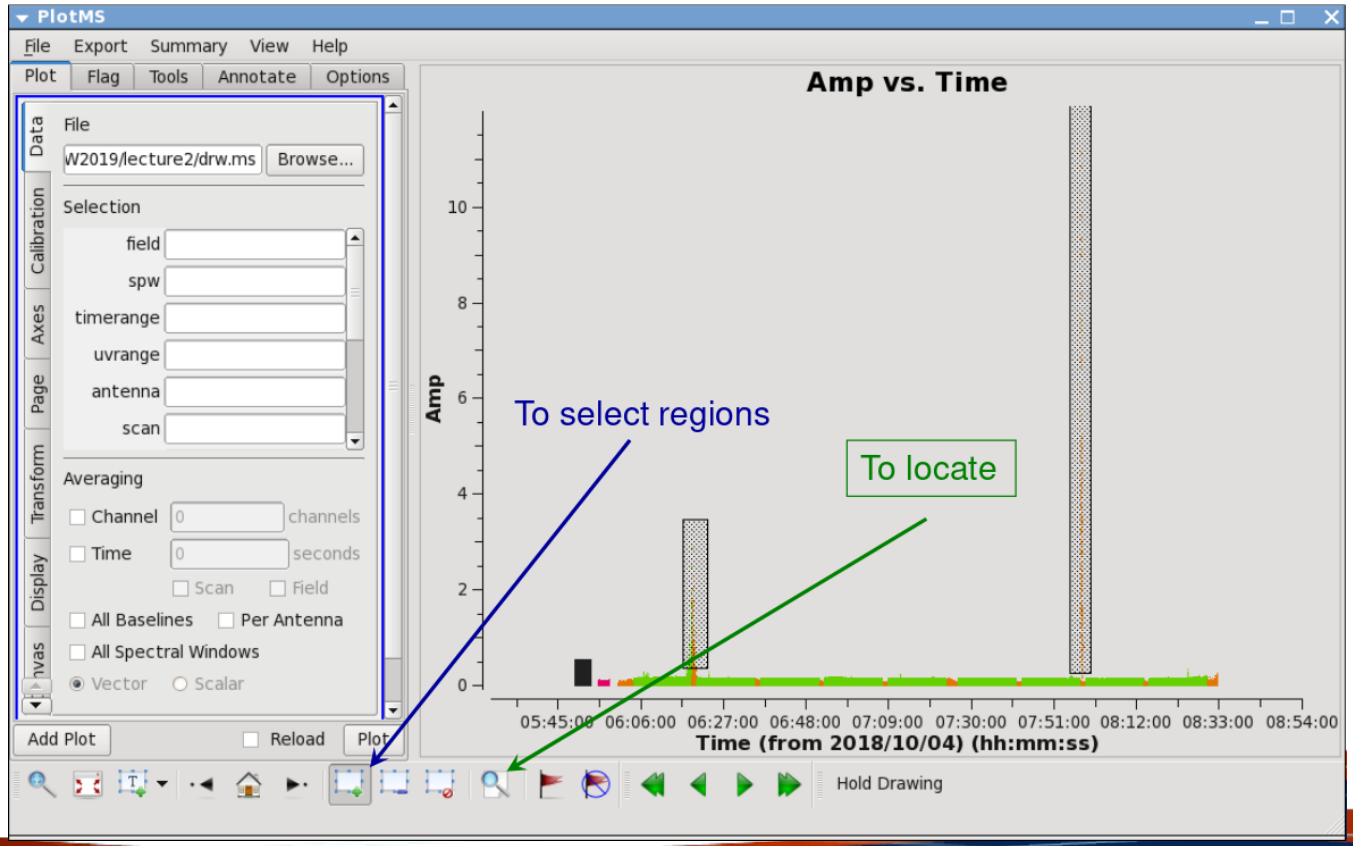
- Flag tool may miss low-level bad data – it only flags in the regions



Example *plotms* editing

Goal: **list outliers with** locate **tool icon**

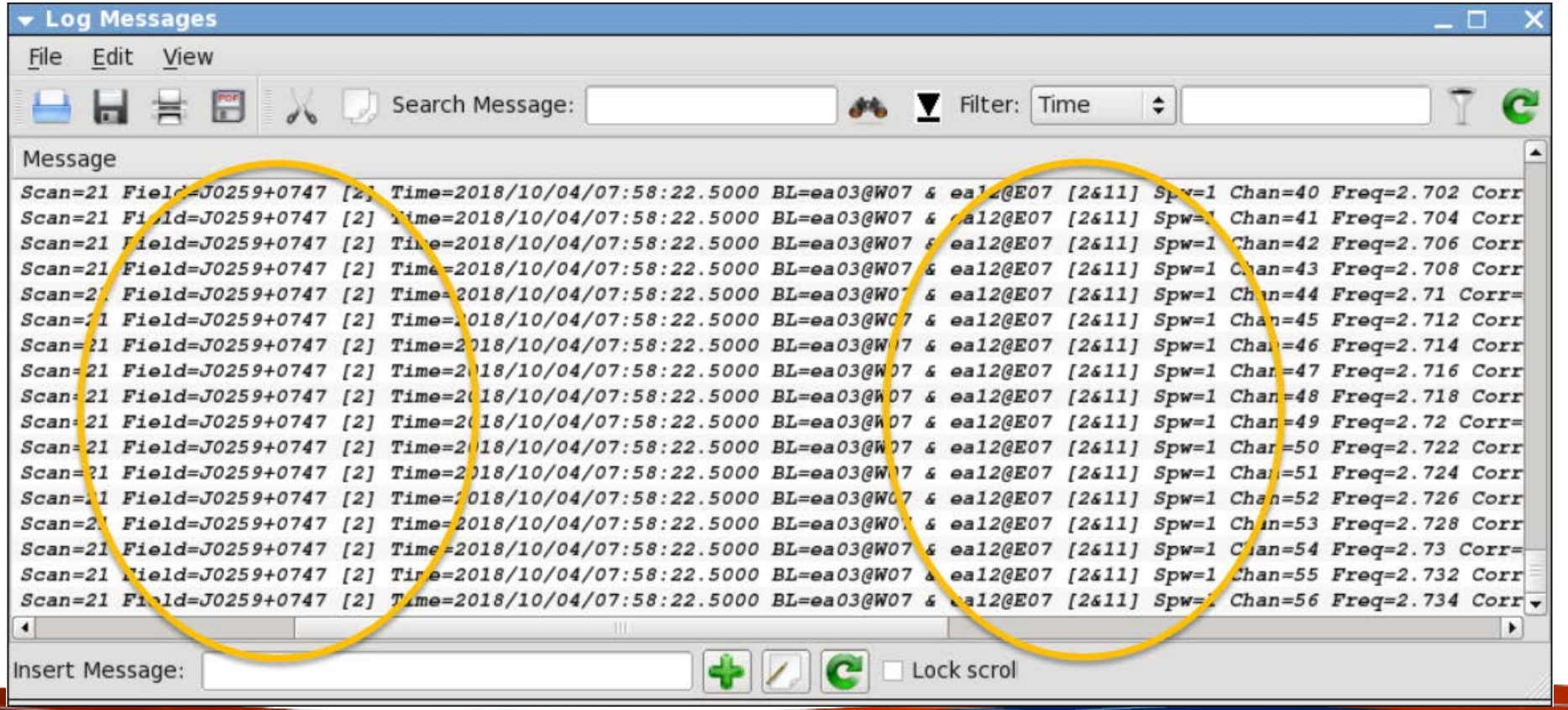
- Locate **will list visibilities in region in message log (can be many!)**



Example *plotms* editing

Goal: identify common identifying parameters with locate

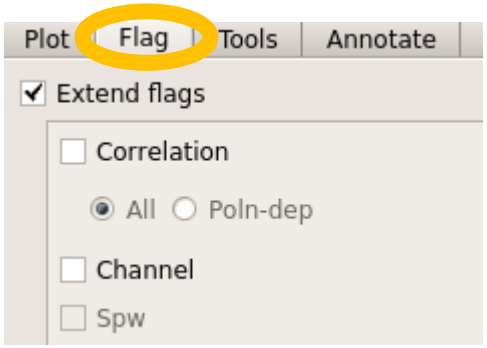
- all channels in spw 1, on baseline ea03-ea12 at 7h58m in scan 21



Important final notes

FYI:

- Be careful when using *plotms* for editing data
  - Interactive unflag possibility may not properly unflag
  - No flag backup file (must use *flagmanager*)
- Editing data with *plotms* perhaps may require extending the flags using the Flag tab:



- Probably best practice:
  1) Use the locate feature and message log window in *plotms* to identify bad data
  2) Flag the bad data using *flagcmd* or *flagdata*
  3) Check that all bad data is gone (extended the flags?)
  4) Repeat if necessary (new bad data may appear)

Finally!(?)

Summary of **visibility data inspection** achievements

- ✓ Data structure is understood, reference antenna is picked
- ✓ Calibrators (flux density, bandpass, gain) are identified
- ✓ Bad antennas and other instrumental flaws are flagged
- ✓ RFI is removed (as much as possible); Hanning smooth?
- ✓ Residual bad individual visibilities are flagged (at least on calibrators – bad target data can be flagged before imaging)

- Maybe inspect (some parts) of the data again to make sure

  More flagging may need to be done during/after calibration steps
- **Ready to start with data calibration** (next lecture)

www.nrao.edu
science.nrao.edu
public.nrao.edu

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