

Data Processing: Visibility Calibration

George J. Bendo
Jodrell Bank Centre for Astrophysics
The University of Manchester



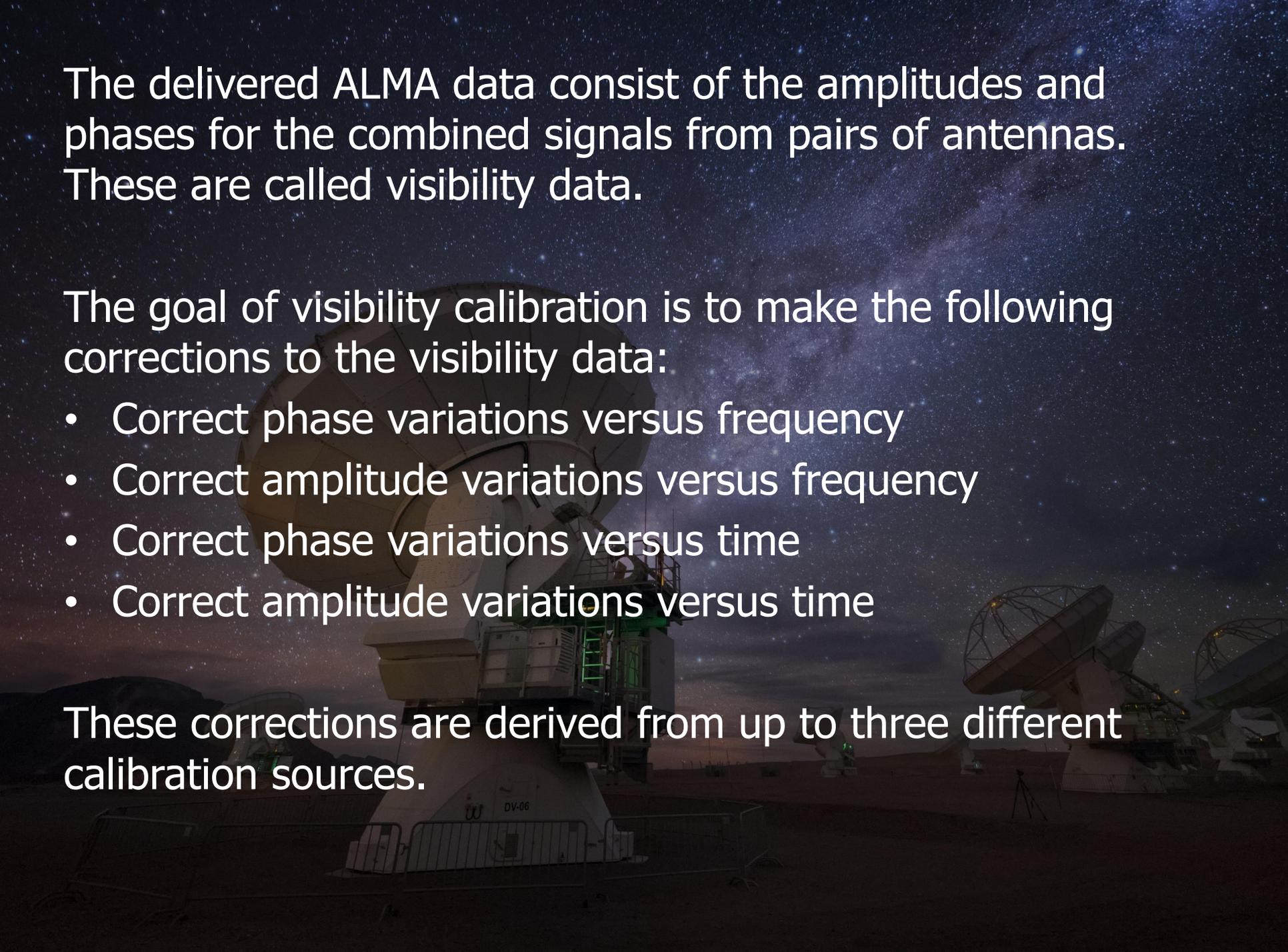
This presentation has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 730562 [RadioNet].

The delivered ALMA data consist of the amplitudes and phases for the combined signals from pairs of antennas. These are called visibility data.

The goal of visibility calibration is to make the following corrections to the visibility data:

- Correct phase variations versus frequency
- Correct amplitude variations versus frequency
- Correct phase variations versus time
- Correct amplitude variations versus time

These corrections are derived from up to three different calibration sources.

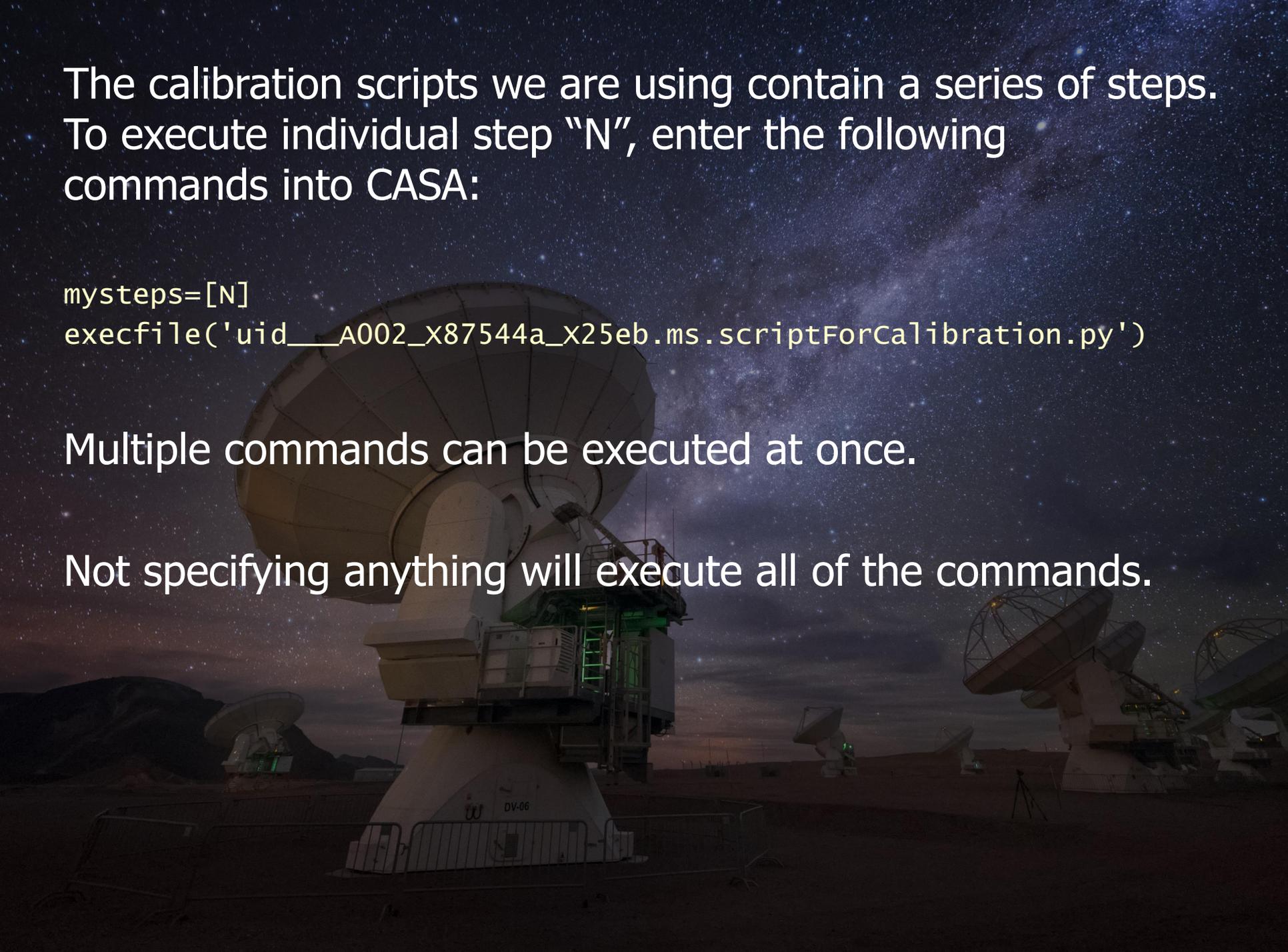


The calibration scripts we are using contain a series of steps. To execute individual step "N", enter the following commands into CASA:

```
mysteps=[N]  
execfile('uid___A002_x87544a_x25eb.ms.scriptForCalibration.py')
```

Multiple commands can be executed at once.

Not specifying anything will execute all of the commands.



The rest of this presentation describes each calibration step in more detail.

The data processing steps can be divided into three broad groups:

0-1: Data import

2-8: A priori calibration

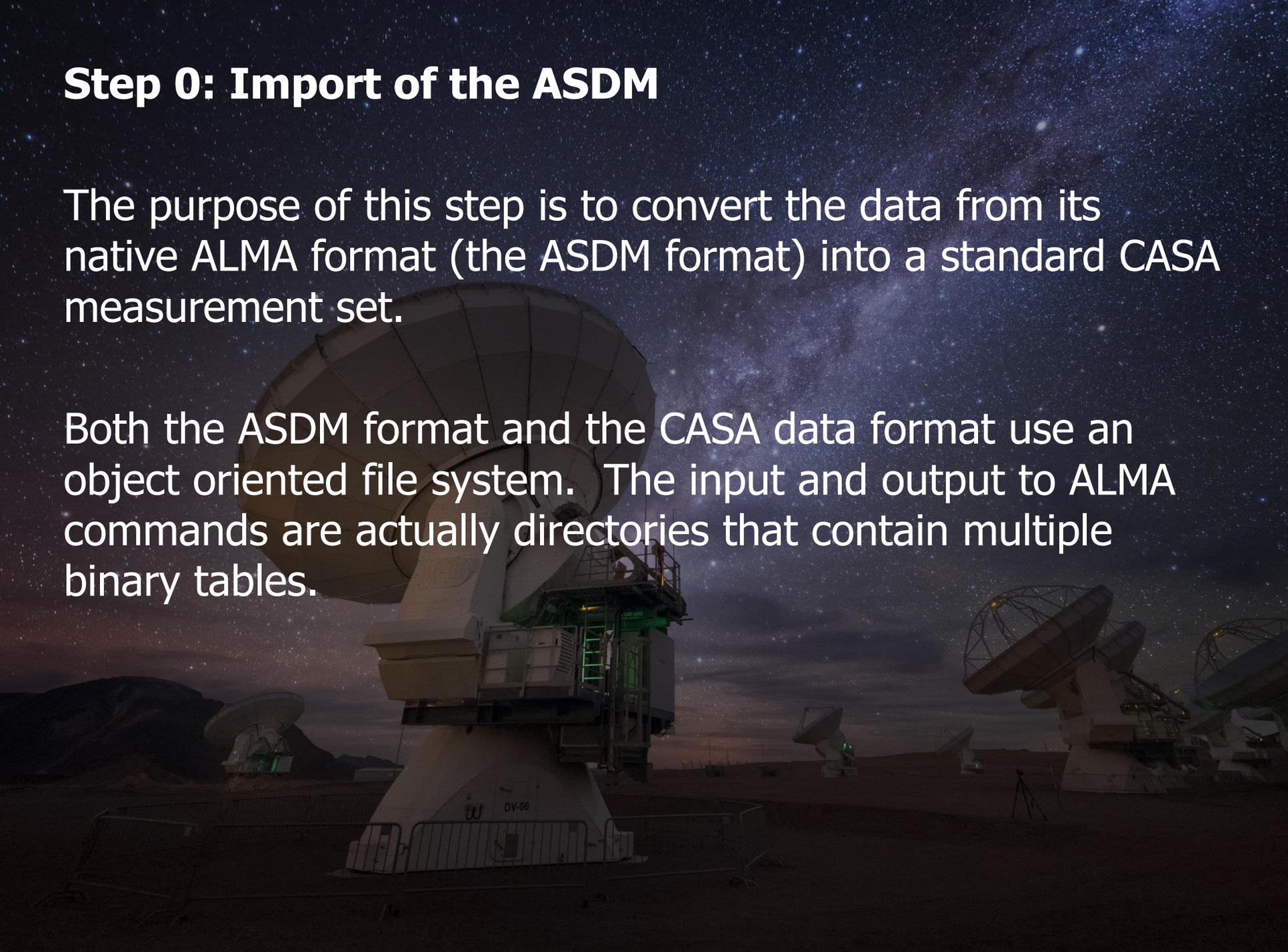
9-19: Calibration



Step 0: Import of the ASDM

The purpose of this step is to convert the data from its native ALMA format (the ASDM format) into a standard CASA measurement set.

Both the ASDM format and the CASA data format use an object oriented file system. The input and output to ALMA commands are actually directories that contain multiple binary tables.



Contents of the example measurement set:

ANTENNA/	POLARIZATION/	table.f17asdminde	table.f24
ASDM_ANTENNA/	PROCESSOR/	table.f18	table.f24_TSM1
ASDM_CALATMOSPHERE/	SOURCE/	table.f19	table.f24_TSM2
ASDM_CALWVR/	SPECTRAL_WINDOW/	table.f2	table.f24_TSM3
ASDM_CORRELATORMODE/	STATE/	table.f20	table.f24_TSM4
ASDM_RECEIVER/	SYSCAL/	table.f20_TSM0	table.f3
ASDM_SBSUMMARY/	SYSPower/	table.f21	table.f4
ASDM_SOURCE/	WEATHER/	table.f21_TSM1	table.f5
ASDM_STATION/	table.dat	table.f21_TSM2	table.f6
CALDEVICE/	table.f1	table.f21_TSM3	table.f7
DATA_DESCRIPTION/	table.f10	table.f21_TSM4	table.f8
FEED/	table.f11	table.f22	table.f9
FIELD/	table.f12	table.f22_TSM1	table.info
FLAG_CMD/	table.f13	table.f22_TSM2	table.lock
HISTORY/	table.f14	table.f23	
OBSERVATION/	table.f15	table.f23_TSM1	
POINTING/	table.f16	table.f23_TSM2	

Step 1: Fix SYSCAL table times

The system temperature (T_{sys}) measurements may include recording errors. Occasionally, multiple measurements have the same timestamp. This step corrects these errors.



Step 2: Fixplanets

This command fixes coordinate information for any Solar System objects in the measurement set (which may have had their coordinates set to 0 during the observations).



Step 3: Listobs

The listobs command will list a large amount of information about the observations. We run this step again later in step 9, so we will look at the listobs file in more detail then.

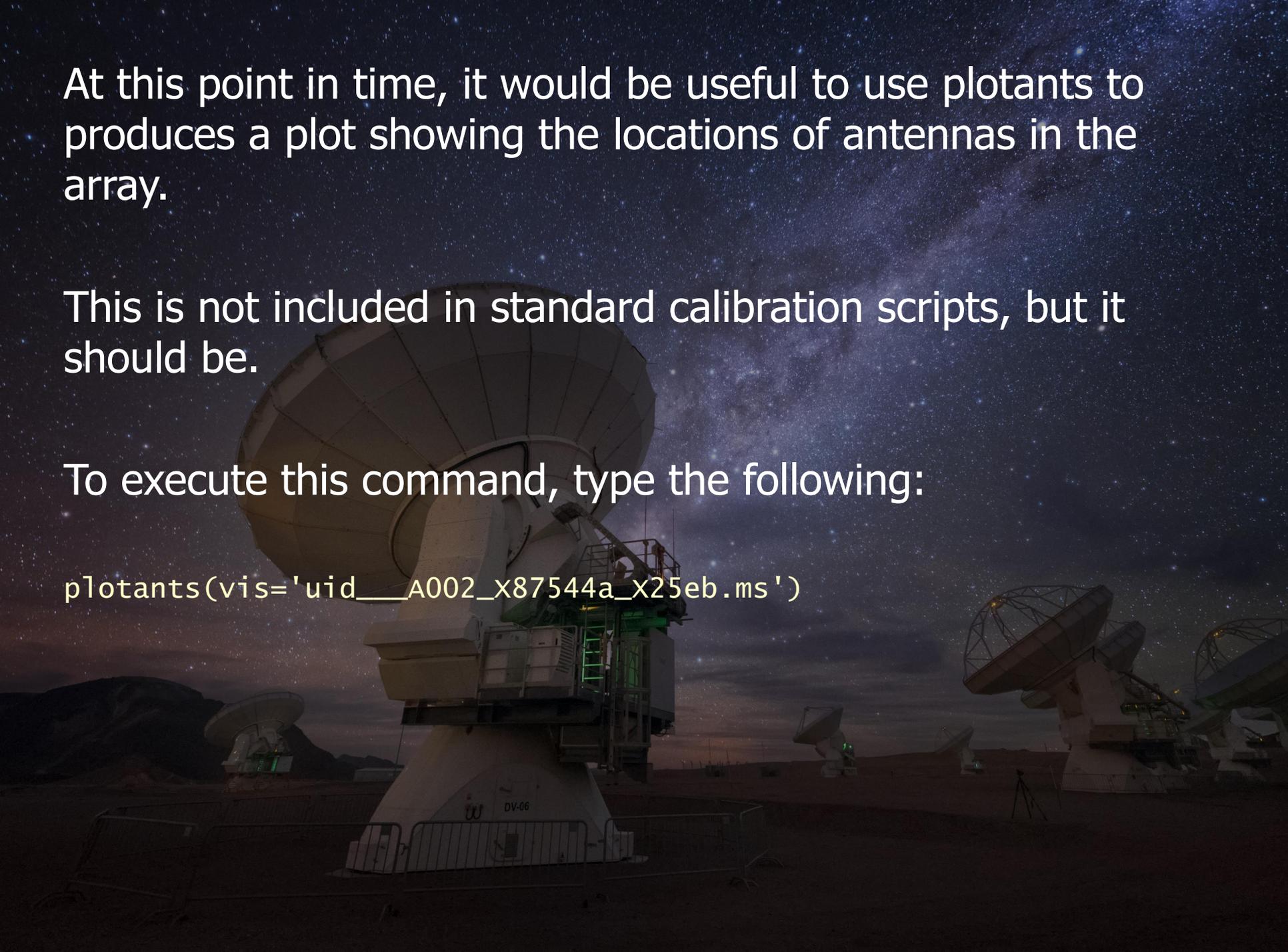


At this point in time, it would be useful to use `plotants` to produce a plot showing the locations of antennas in the array.

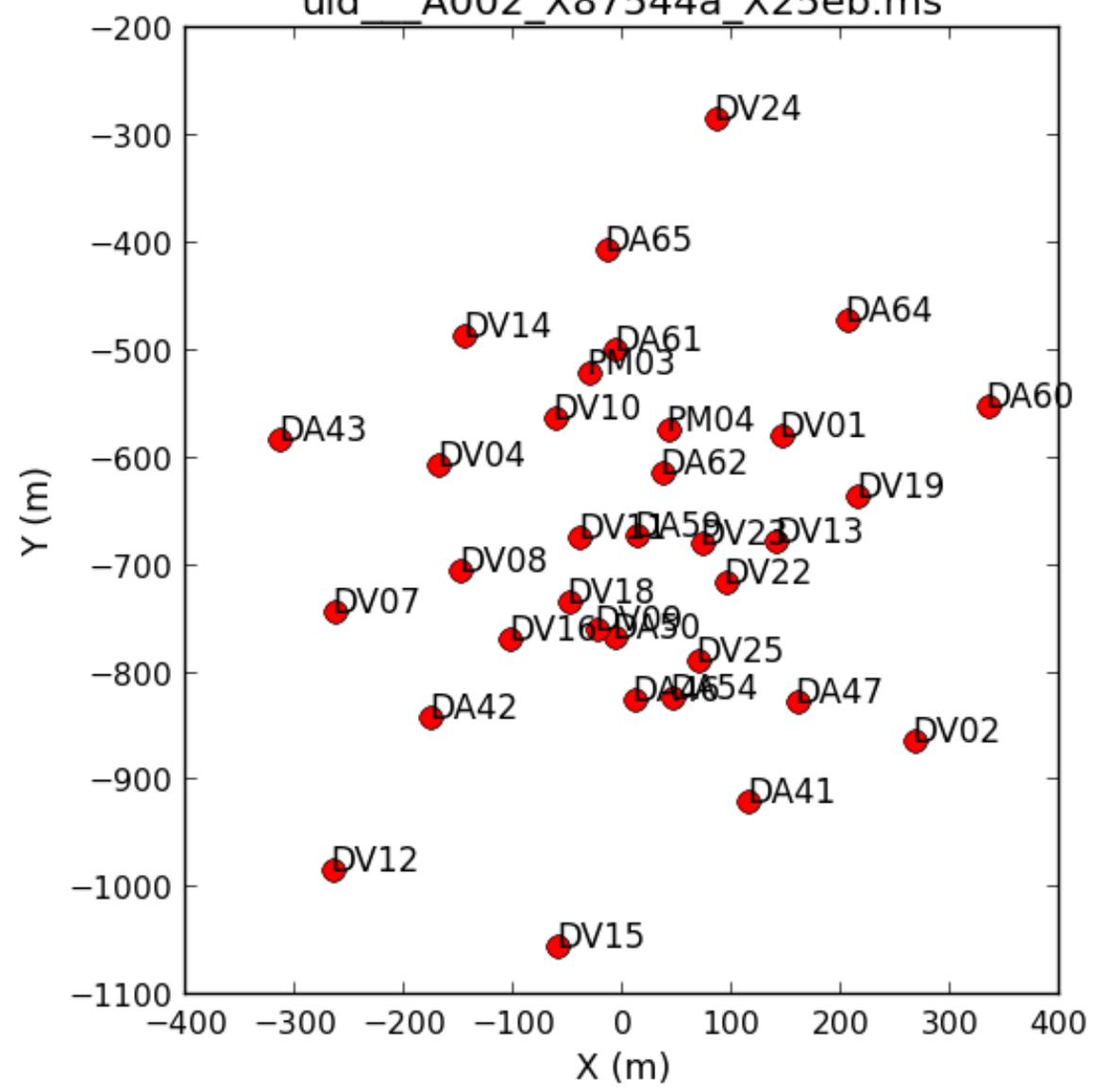
This is not included in standard calibration scripts, but it should be.

To execute this command, type the following:

```
plotants(vis='uid___A002_x87544a_x25eb.ms')
```

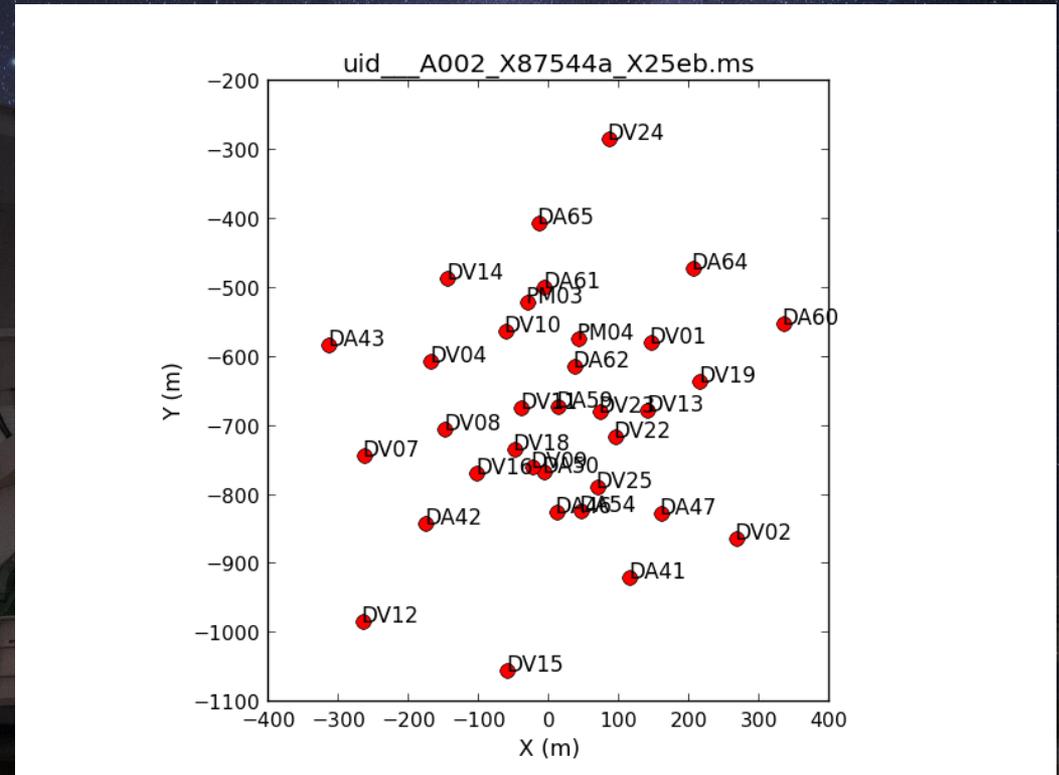


uid __ A002_X87544a_X25eb.ms



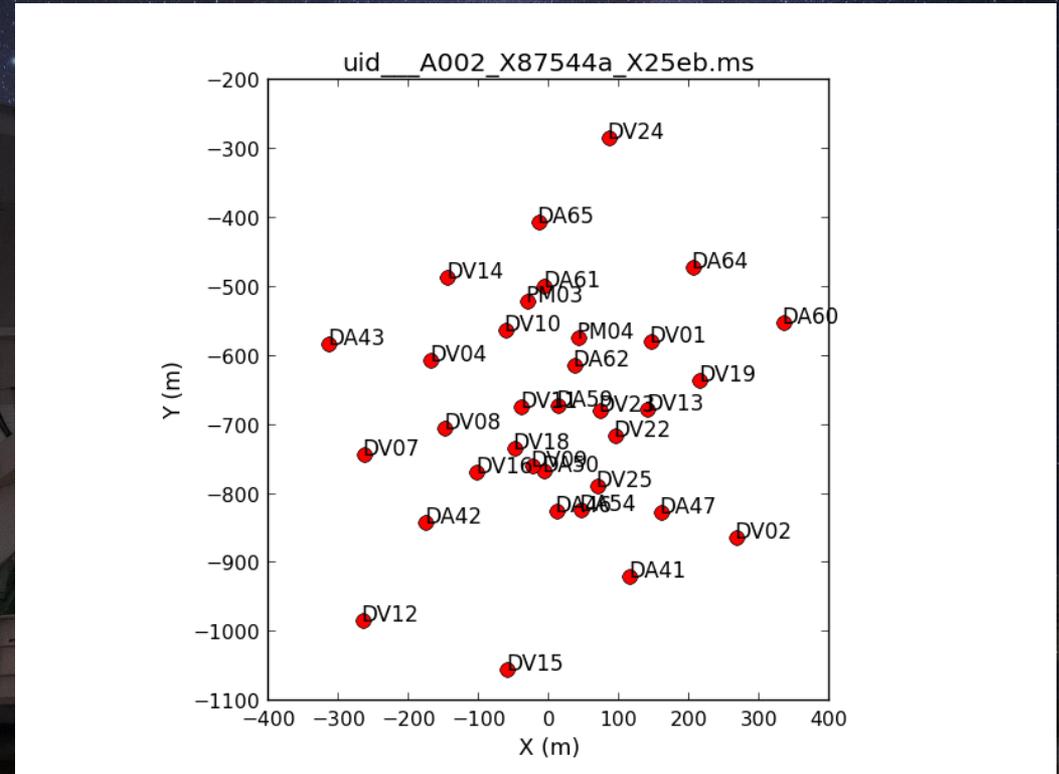
This plot can be used to select the reference antenna, which is used in many of the data processing steps.

The antenna needs to be from the centre of the array.



Antennas on long baselines will tend to produce noisier data than antennas on short baselines.

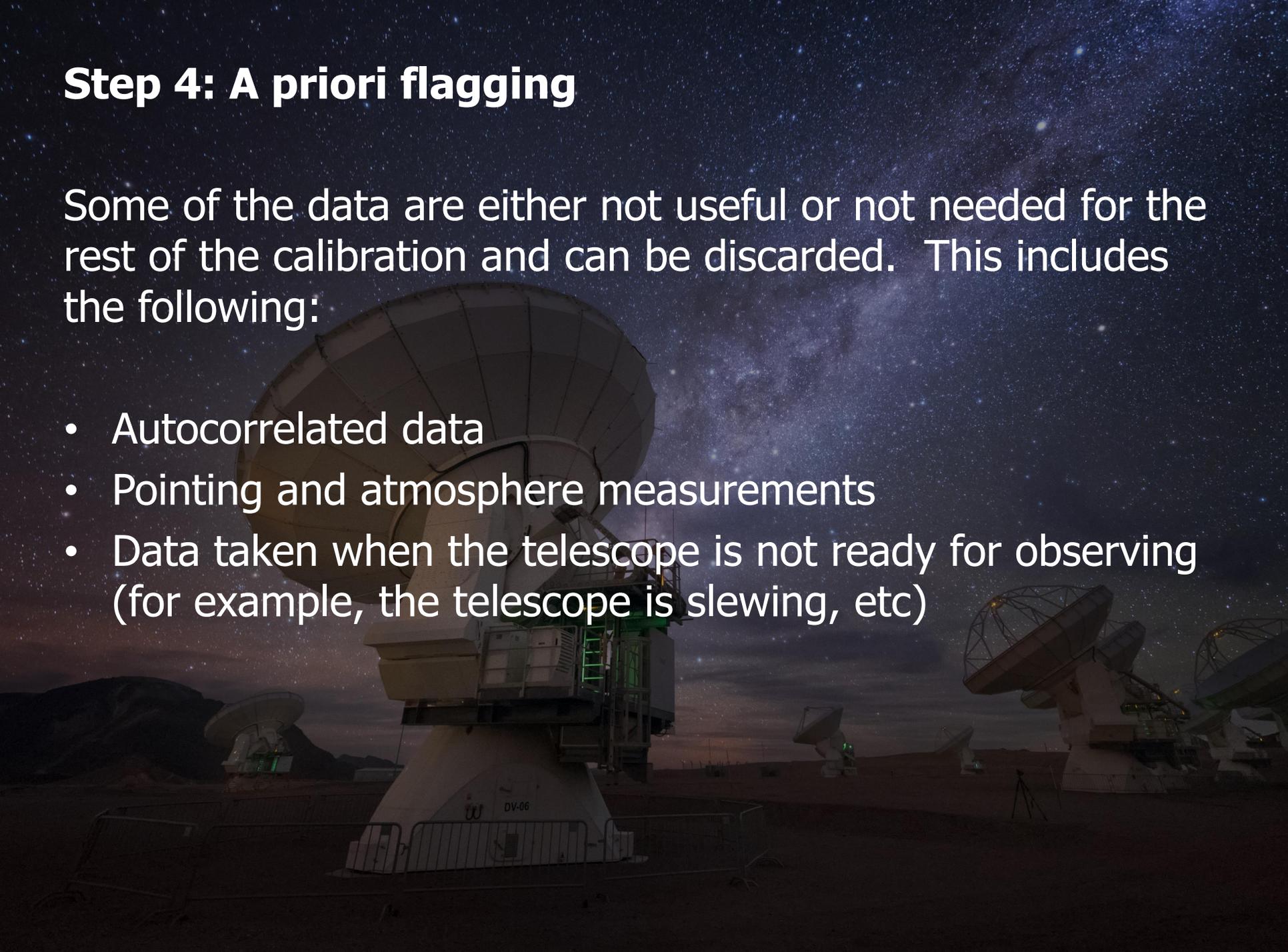
This figure can be used to diagnose these types of issues.



Step 4: A priori flagging

Some of the data are either not useful or not needed for the rest of the calibration and can be discarded. This includes the following:

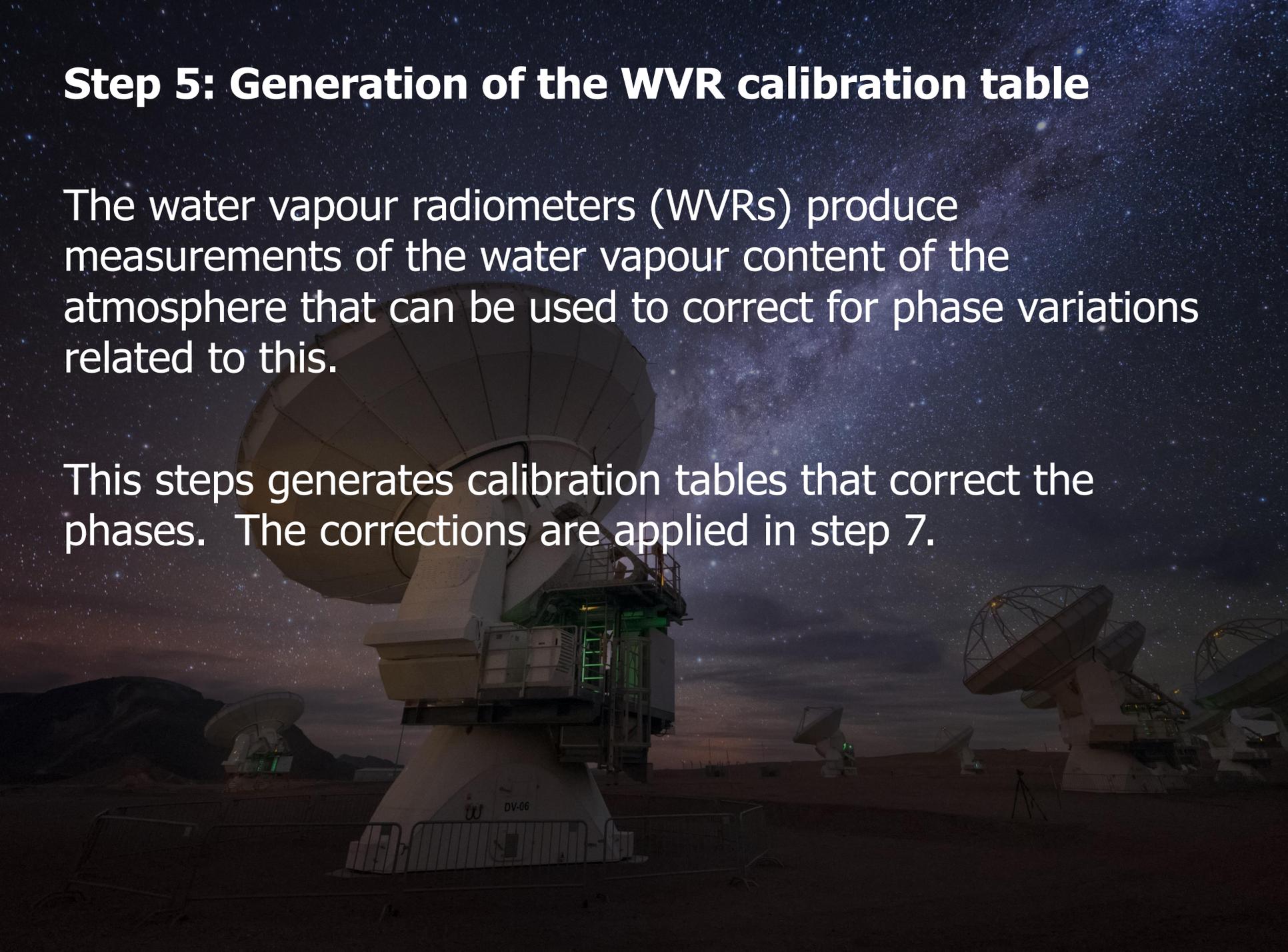
- Autocorrelated data
- Pointing and atmosphere measurements
- Data taken when the telescope is not ready for observing (for example, the telescope is slewing, etc)



Step 5: Generation of the WVR calibration table

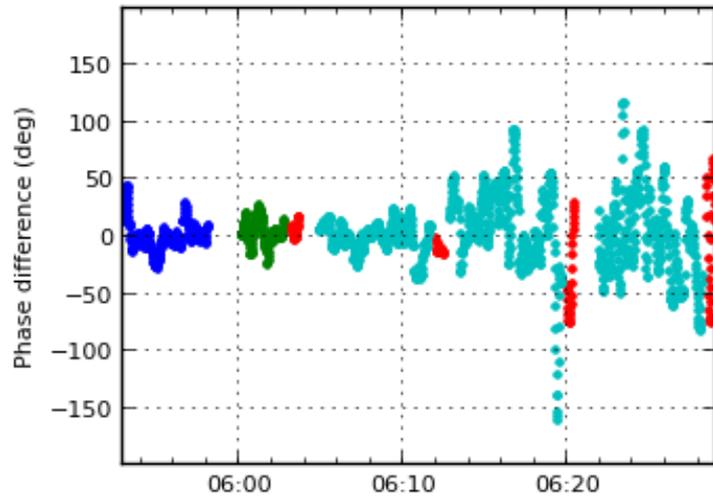
The water vapour radiometers (WVRs) produce measurements of the water vapour content of the atmosphere that can be used to correct for phase variations related to this.

This step generates calibration tables that correct the phases. The corrections are applied in step 7.

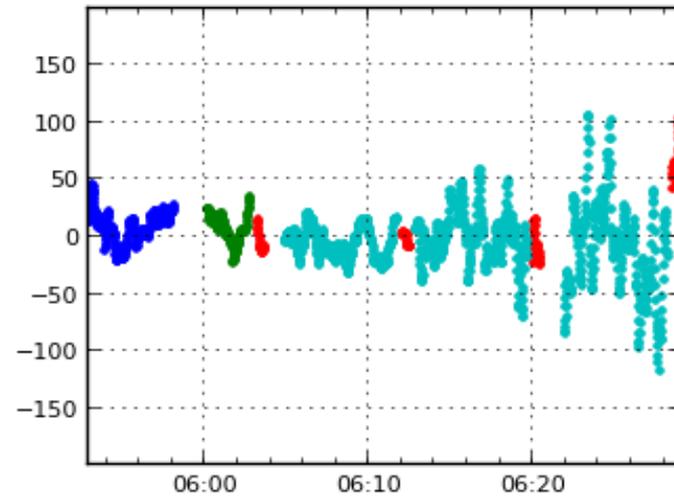


uid__A002_X87544a_X25eb.ms.wvr computed for uid__A002_X87544a_X25eb.ms

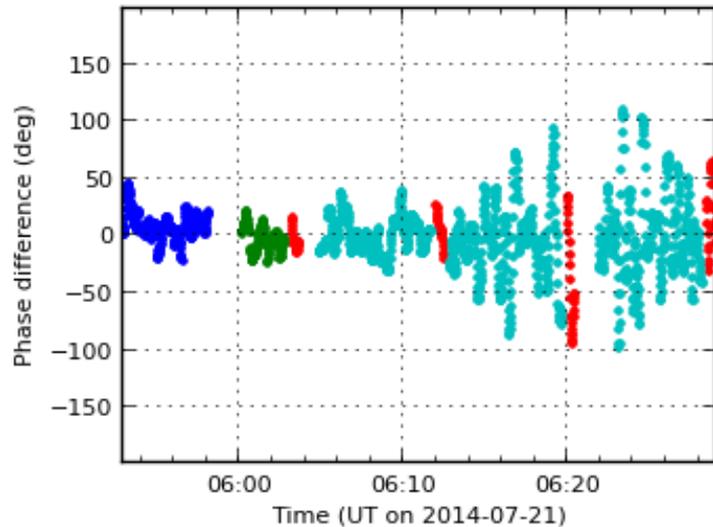
290m Baseline 0-20=DA41-DV11, spw17=99.4GHz



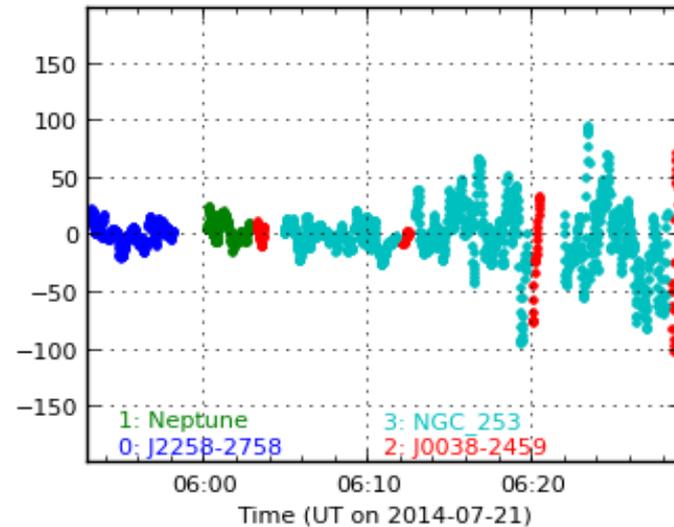
216m Baseline 1-20=DA42-DV11, spw17=99.4GHz



290m Baseline 2-20=DA43-DV11, spw17=99.4GHz



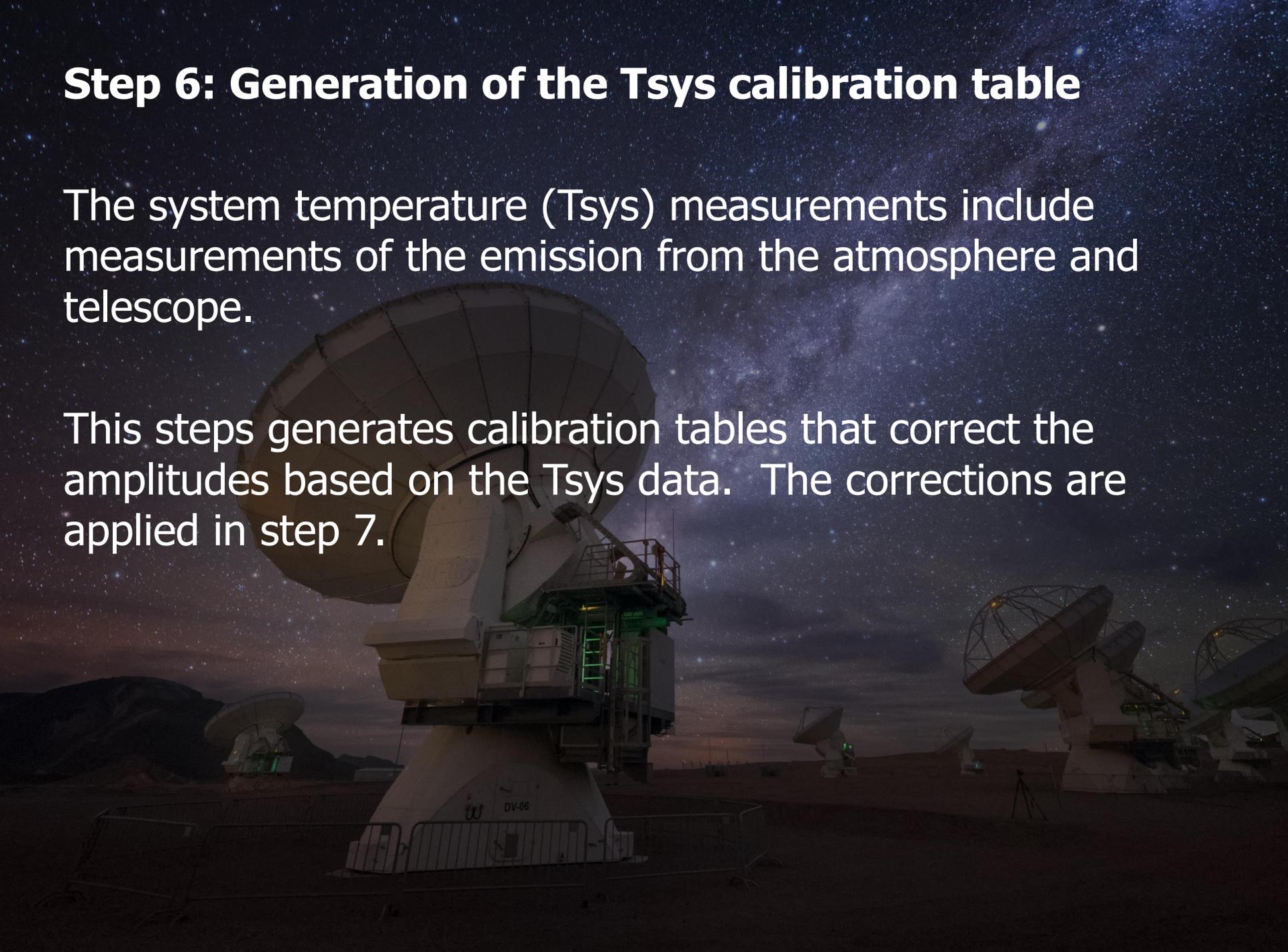
160m Baseline 3-20=DA46-DV11, spw17=99.4GHz



Step 6: Generation of the Tsys calibration table

The system temperature (T_{sys}) measurements include measurements of the emission from the atmosphere and telescope.

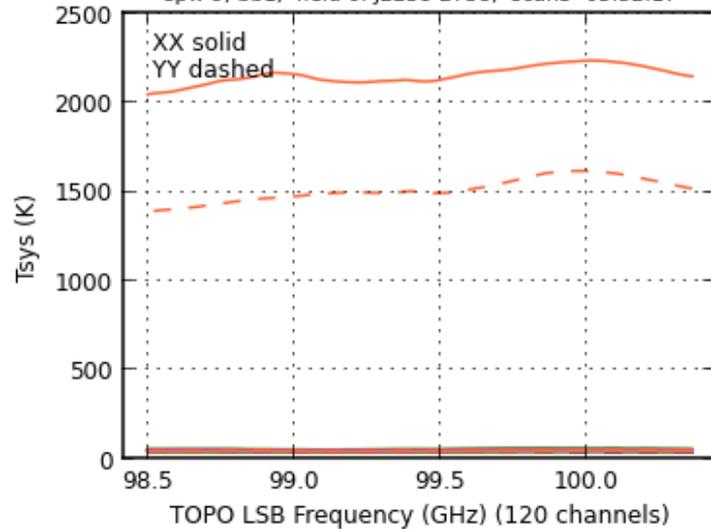
This step generates calibration tables that correct the amplitudes based on the T_{sys} data. The corrections are applied in step 7.



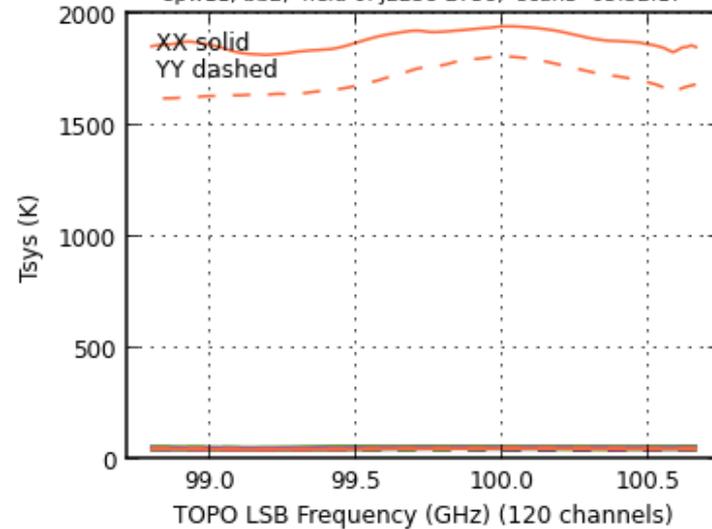
uid__A002_X87544a_X25eb.ms.tsys

DA41 DA42 DA43 DA46 DA47 DA50 DA54 DA59 DA60 DA61 DA62 DA64 DA65 DV01 DV02 DV04 DV07 DV08

spw 9, bb1, field 0: J2258-2758, scan3 05:52:17

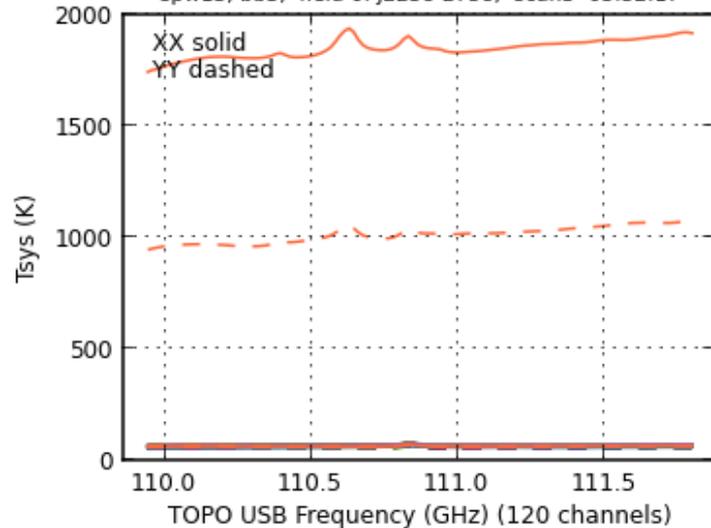


spw11, bb2, field 0: J2258-2758, scan3 05:52:17

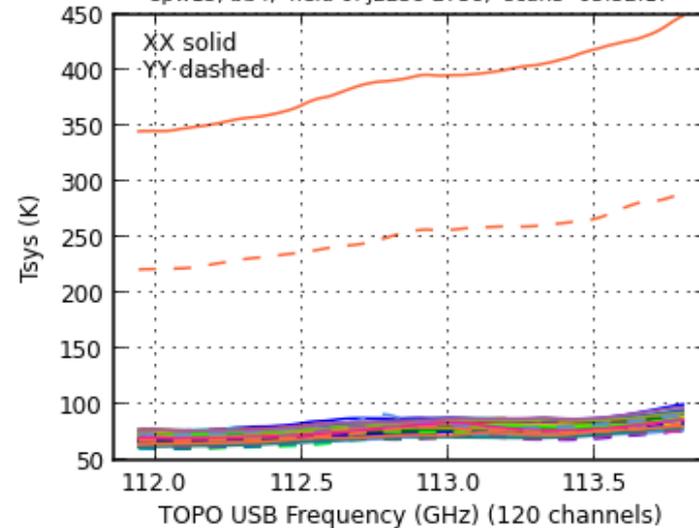


DV09
DV10
DV11
DV12
DV13
DV14
DV15
DV16
DV18
DV19
DV22
DV23
DV24
DV25
PM03
PM04

spw13, bb3, field 0: J2258-2758, scan3 05:52:17



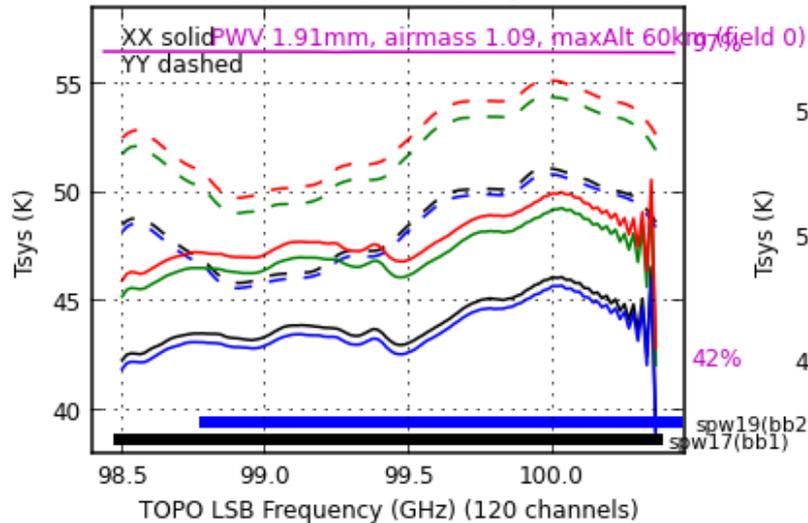
spw15, bb4, field 0: J2258-2758, scan3 05:52:17



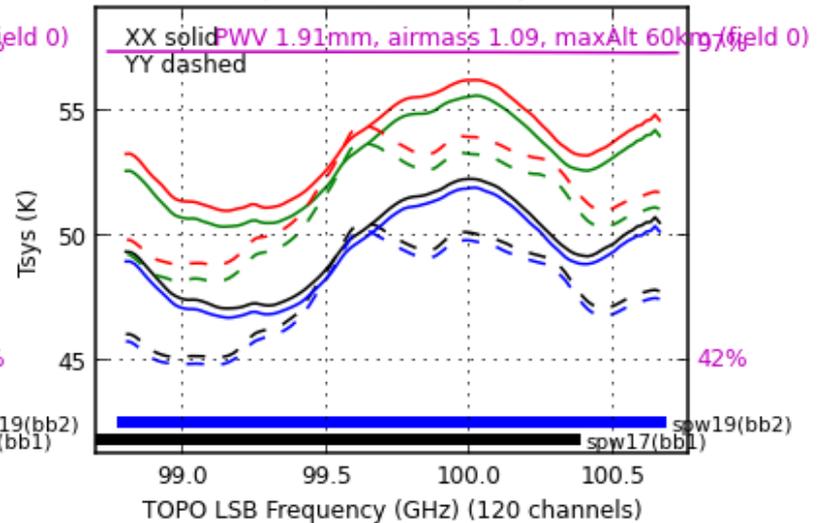
uid__A002_X87544a_X25eb.ms.tsys

UT 05:52:17 05:59:18 06:04:01 06:21:04

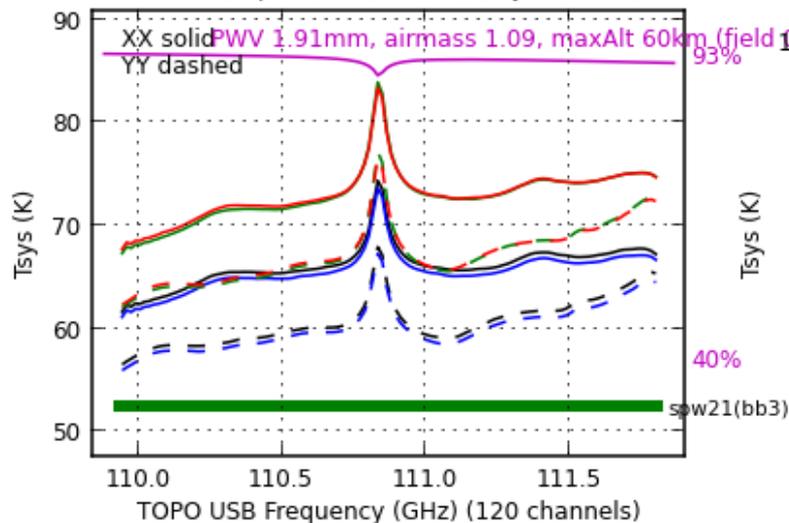
Ant 0: DA41, spw 9, bb1, fields 0,1,3: J2258-2758,N...



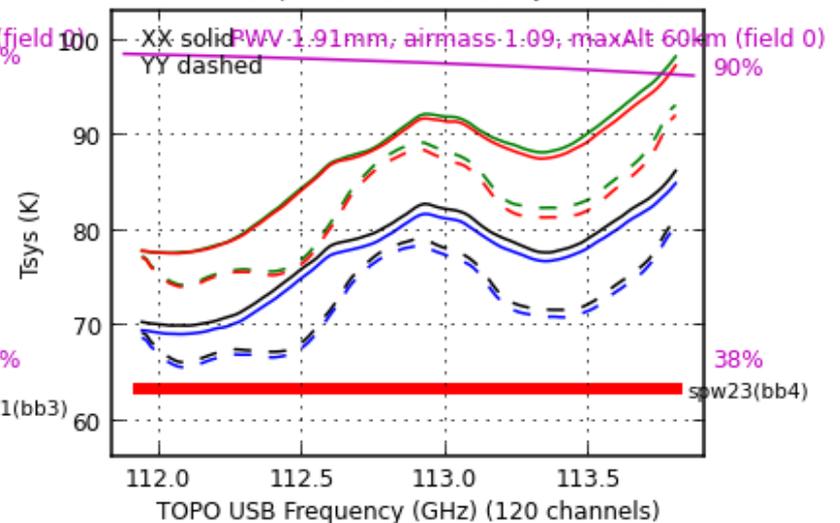
Ant 0: DA41, spw11, bb2, fields 0,1,3: J2258-2758,N...



Ant 0: DA41, spw13, bb3, fields 0,1,3: J2258-2758,N...



Ant 0: DA41, spw15, bb4, fields 0,1,3: J2258-2758,N...



[Generation of the antenna position calibration table]

Corrections to antenna positions can be very important if the antennas are on long baselines. This step generates a calibration table that corrects the positions.

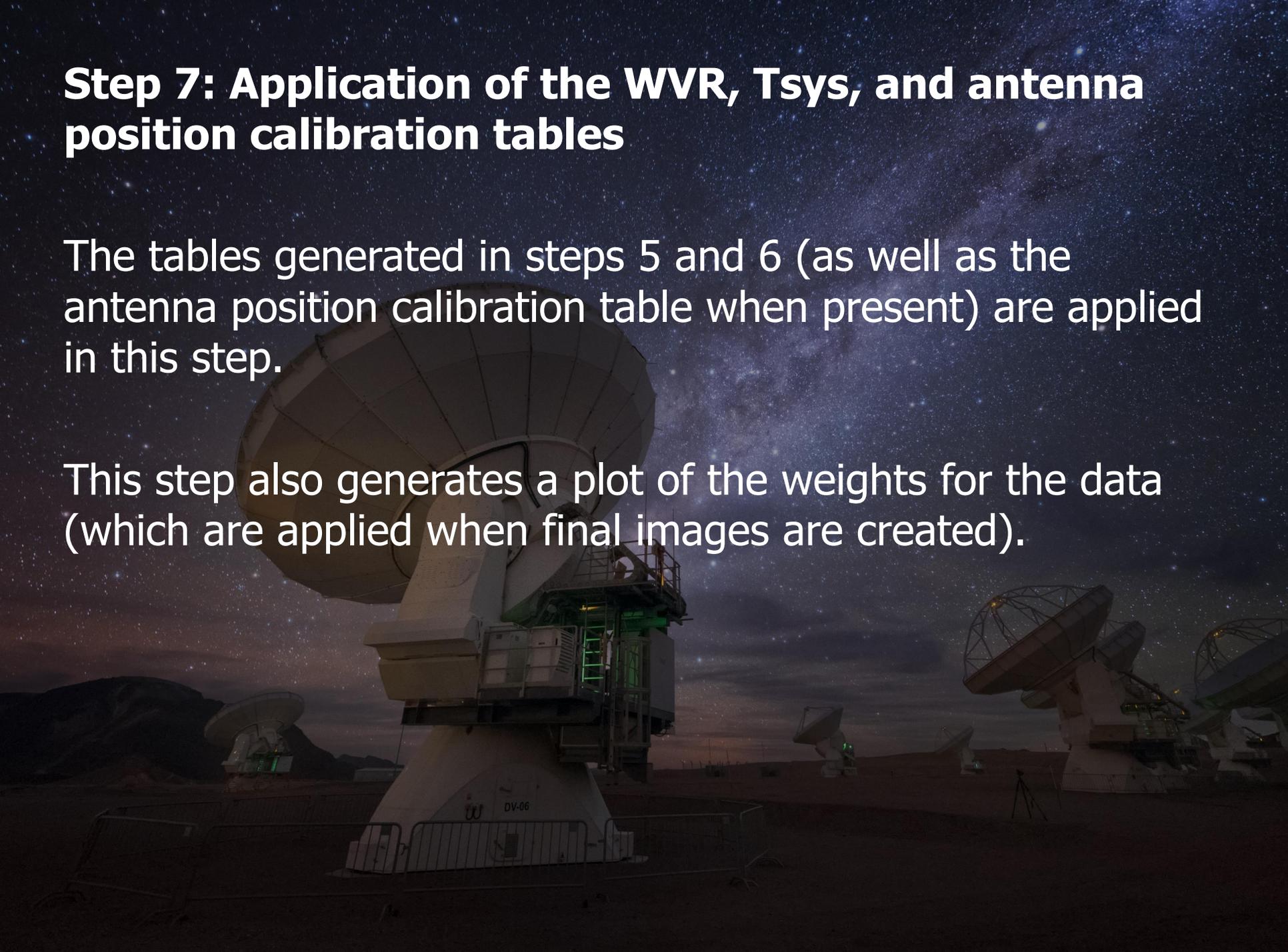
For most other data, though, the corrections may be unnecessary. In this situation, a table filled with zeroes is created.

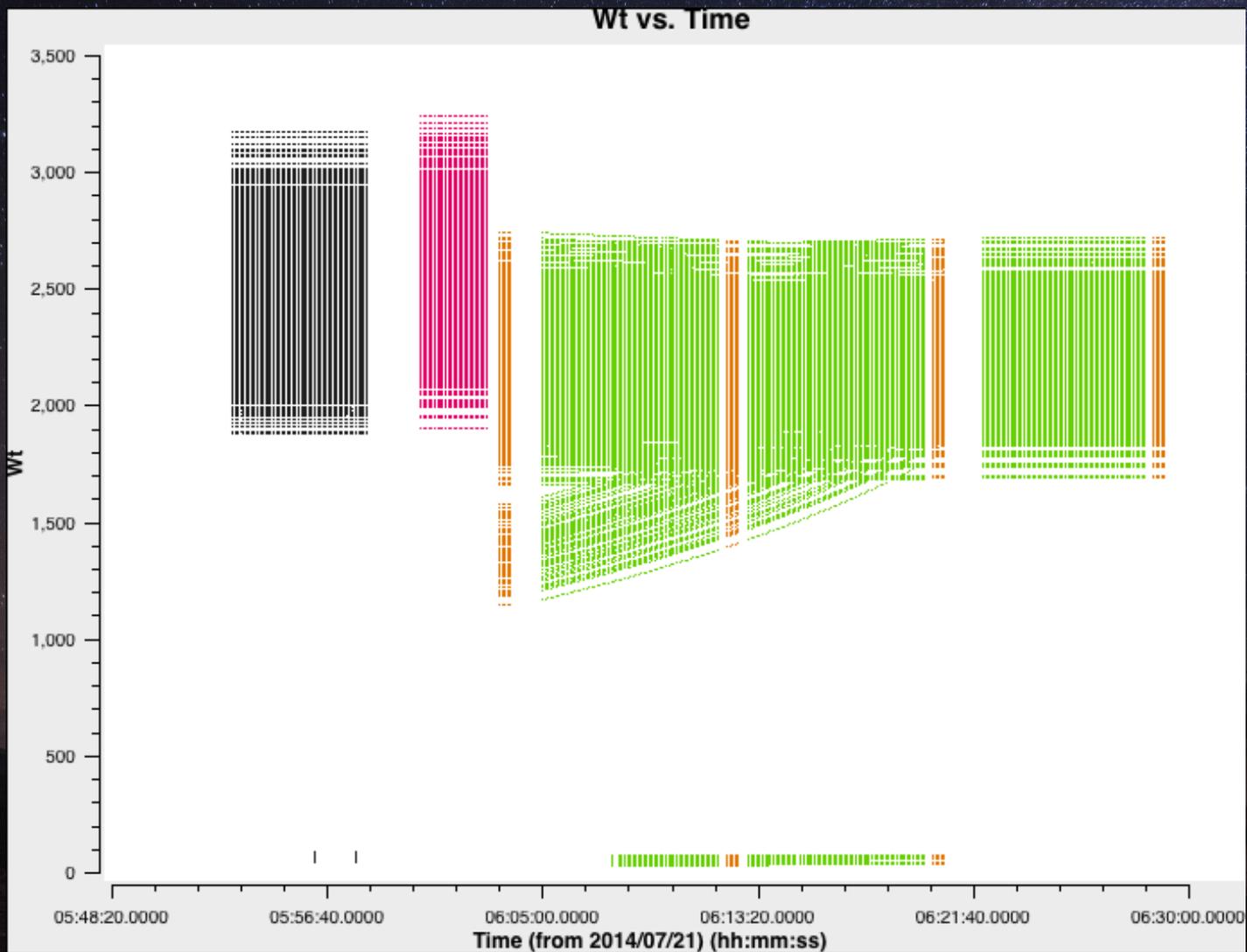
(Depending on when the data are generated, the corrections may have already been incorporated into the measurement set. The script generator did not even create a line for this step, so presumably it is not needed.)

Step 7: Application of the WVR, T_{sys} , and antenna position calibration tables

The tables generated in steps 5 and 6 (as well as the antenna position calibration table when present) are applied in this step.

This step also generates a plot of the weights for the data (which are applied when final images are created).



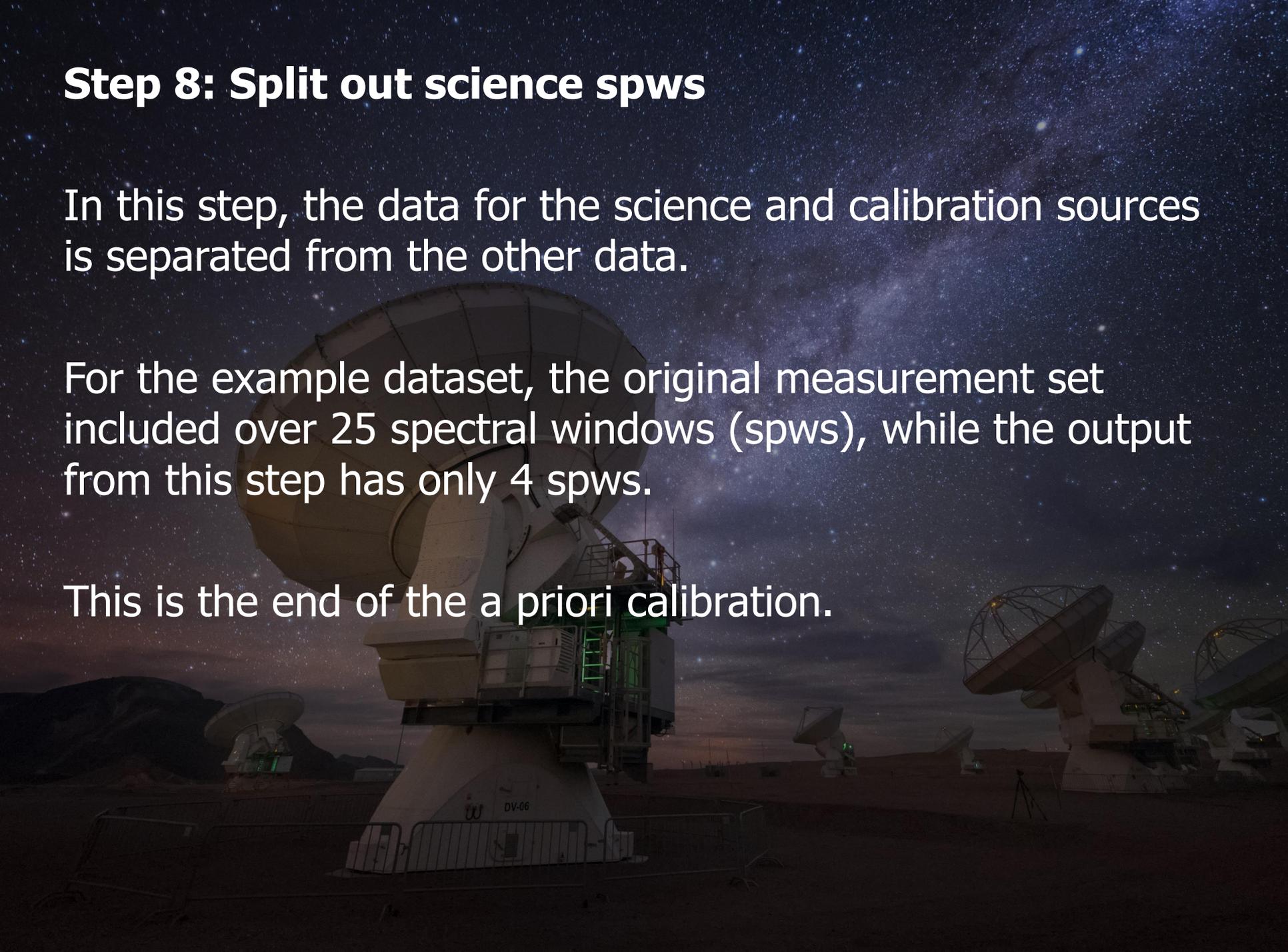


Step 8: Split out science spws

In this step, the data for the science and calibration sources is separated from the other data.

For the example dataset, the original measurement set included over 25 spectral windows (spws), while the output from this step has only 4 spws.

This is the end of the a priori calibration.



Step 9: Listobs and save initial flags

The listobs command will list a large amount of information about the observations, including the following:

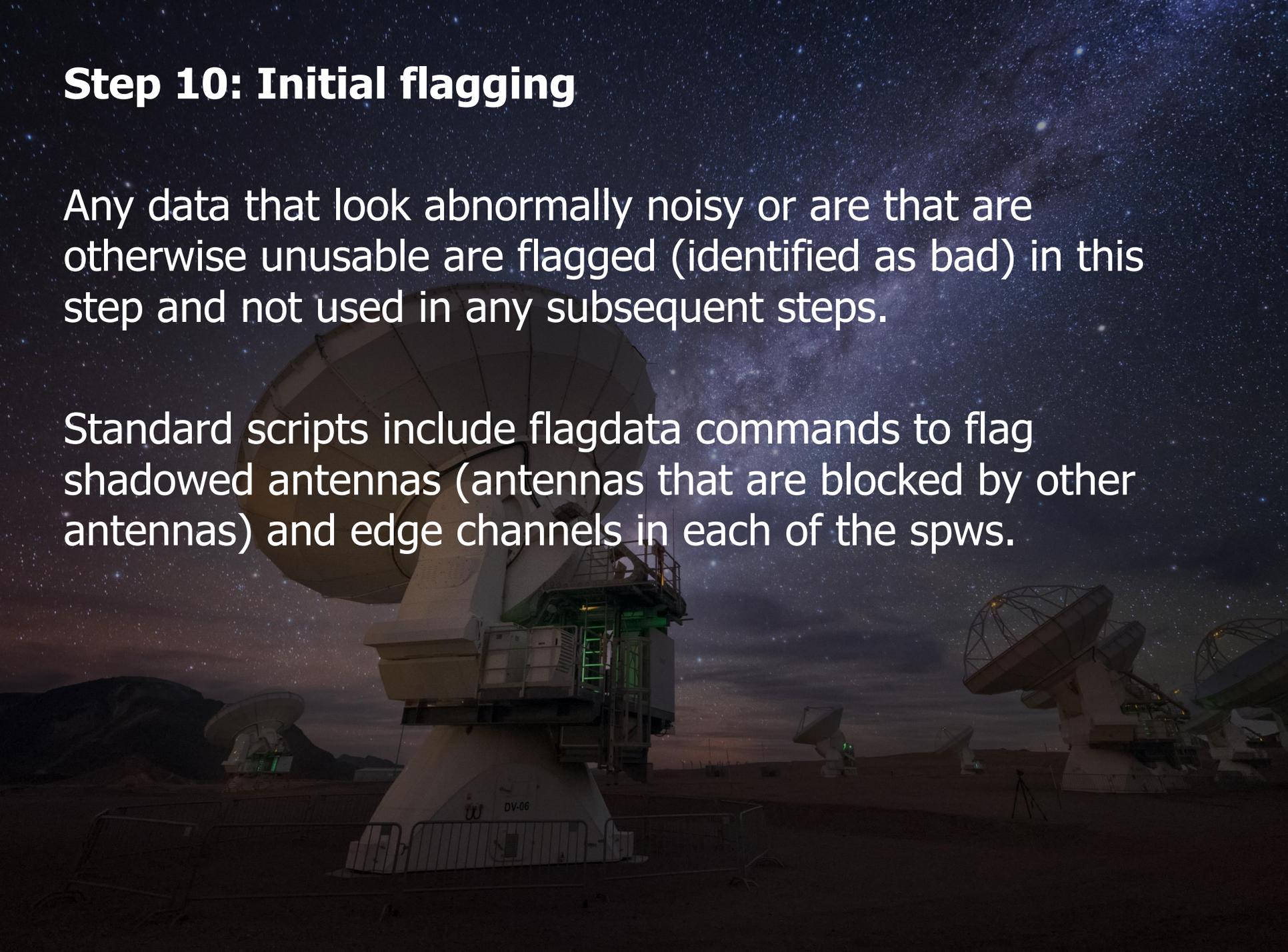
- Basic observation information (PI, project ID, dates, etc)
- List of the sequence of observations performed
- List of the fields
- List of the spectral windows
- List of the sources (the spectral windows for each field)
- List of the antennas

The other part of this step saves data related to flagging. This is done several times throughout the script.

Step 10: Initial flagging

Any data that look abnormally noisy or are that are otherwise unusable are flagged (identified as bad) in this step and not used in any subsequent steps.

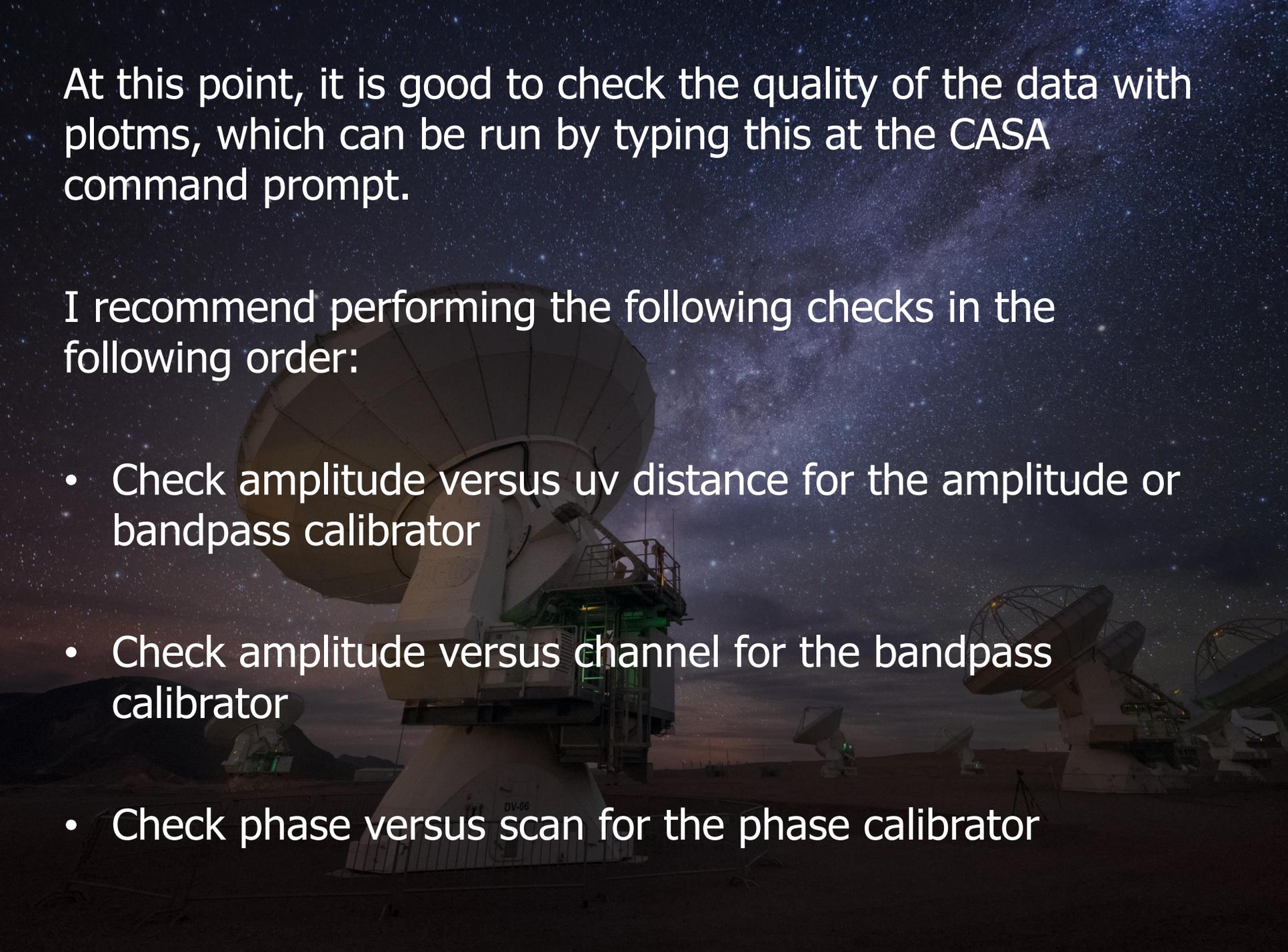
Standard scripts include flagdata commands to flag shadowed antennas (antennas that are blocked by other antennas) and edge channels in each of the spws.



At this point, it is good to check the quality of the data with `plotms`, which can be run by typing this at the CASA command prompt.

I recommend performing the following checks in the following order:

- Check amplitude versus uv distance for the amplitude or bandpass calibrator
- Check amplitude versus channel for the bandpass calibrator
- Check phase versus scan for the phase calibrator



File

Browse...

Data

Selection

field
spw
timerange
uvrange
antenna
scan
corr
array
observation
intent
feed
msselect

Display

Averaging

Channel channels
 Time seconds
 Scan Field
 All Baselines Per Antenna
 All Spectral Windows
 Vector Scalar

Canvas

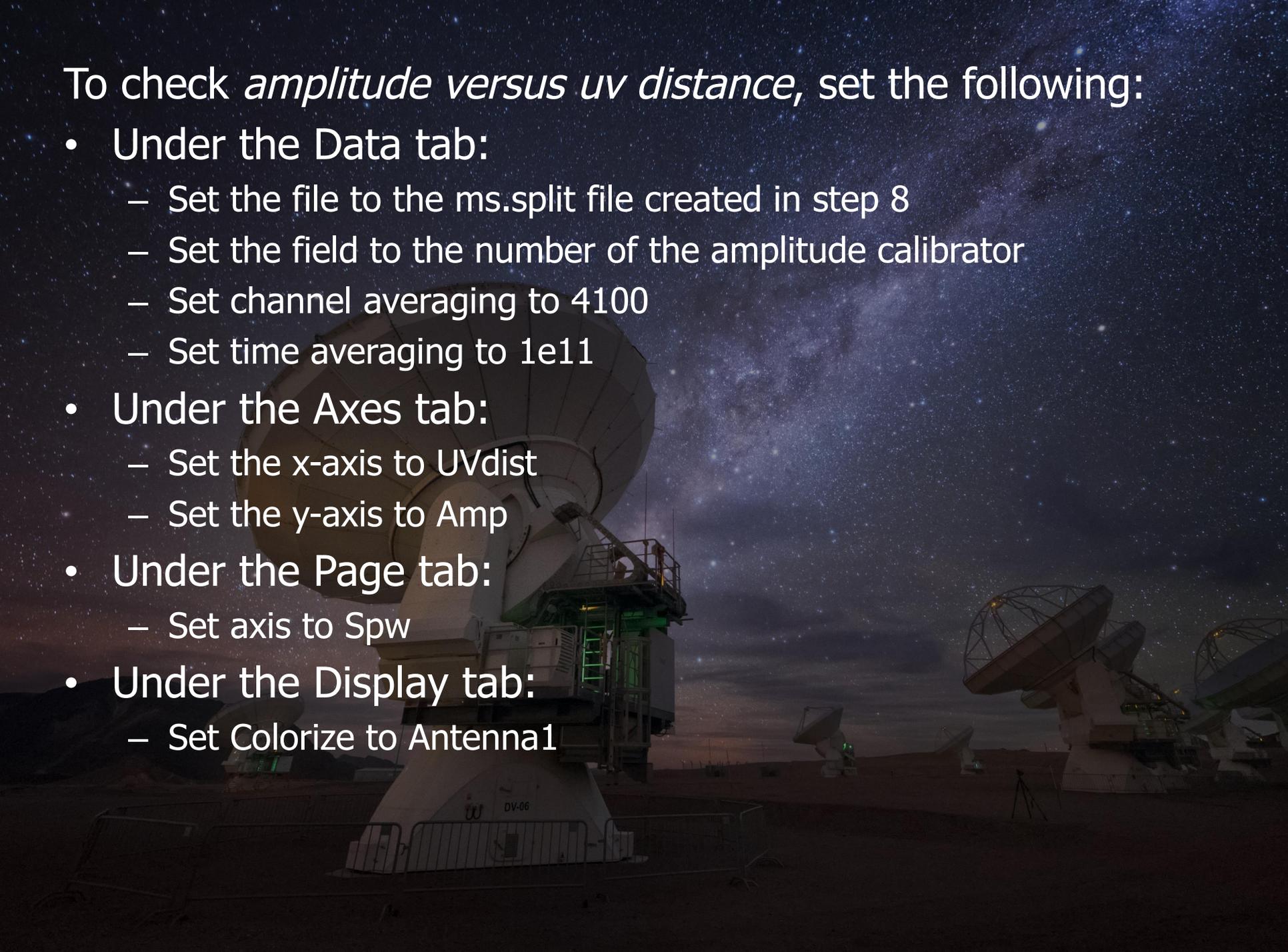
Add Plot

Reload Plot



To check *amplitude versus uv distance*, set the following:

- Under the Data tab:
 - Set the file to the ms.split file created in step 8
 - Set the field to the number of the amplitude calibrator
 - Set channel averaging to 4100
 - Set time averaging to 1e11
- Under the Axes tab:
 - Set the x-axis to UVdist
 - Set the y-axis to Amp
- Under the Page tab:
 - Set axis to Spw
- Under the Display tab:
 - Set Colorize to Antenna1



Plot Flag Tools Annotate Options

File

544a_X25eb.ms.split

Browse...

Data

Selection

field

1

spw

timerange

uvrange

antenna

scan

corr

array

observation

intent

feed

msselect

Calibration

Axes

Page

Transform

Display

Canvas

Averaging

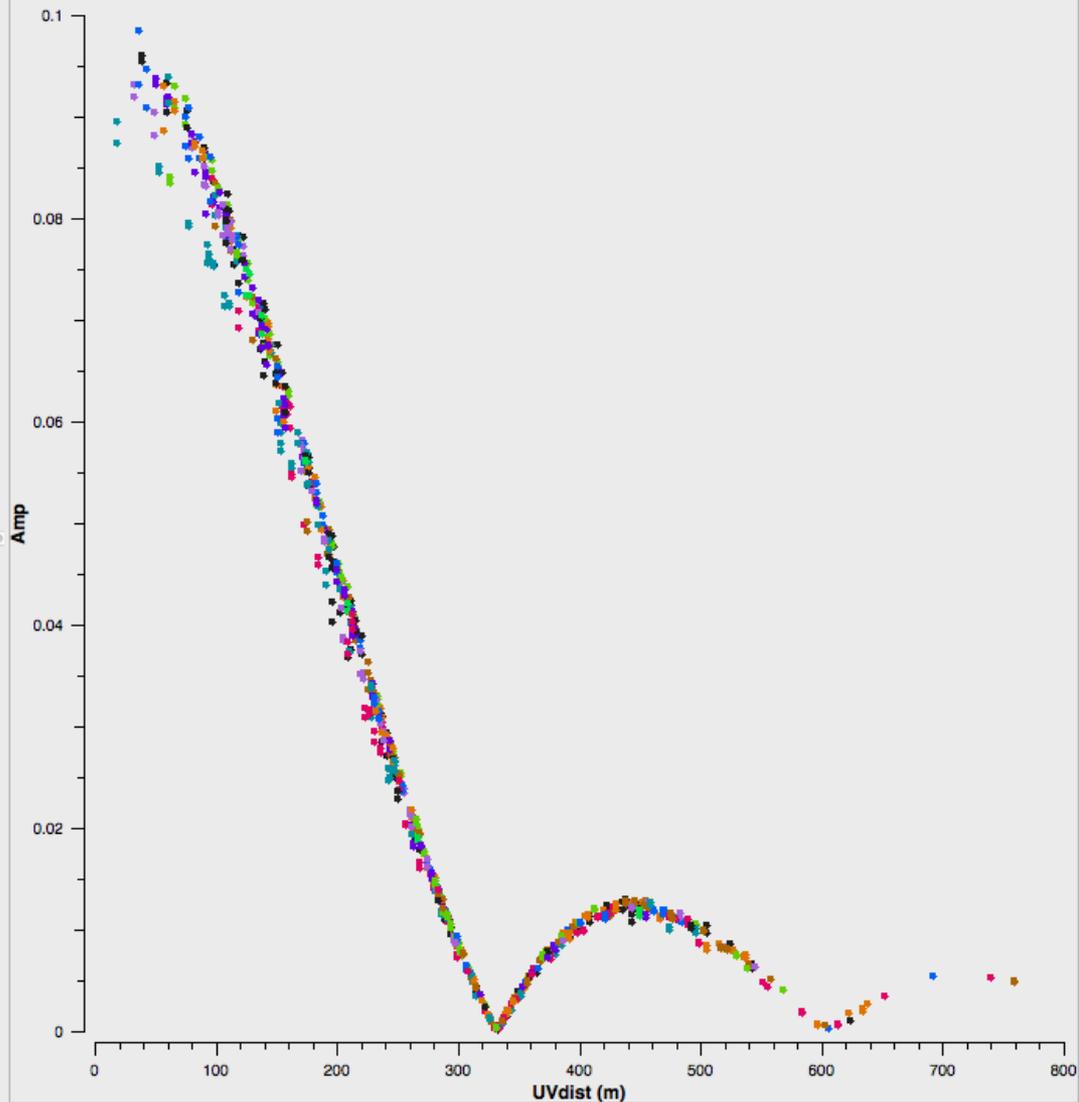
 Channel 4100 channels Time 1e+11 seconds Scan Field All Baselines Per Antenna All Spectral Windows Vector Scalar

Add Plot

Reload

Plot

Amp vs. UVdist Spw: 0



Hold Drawing

X Axis: UVdist

Cached: Attach: Bottom Top

Range:

 Automatic

0

to

 0

Data: Amp

Data Column: data

Cached: Attach: Left Right

Range:

 Automatic

0

to

 0

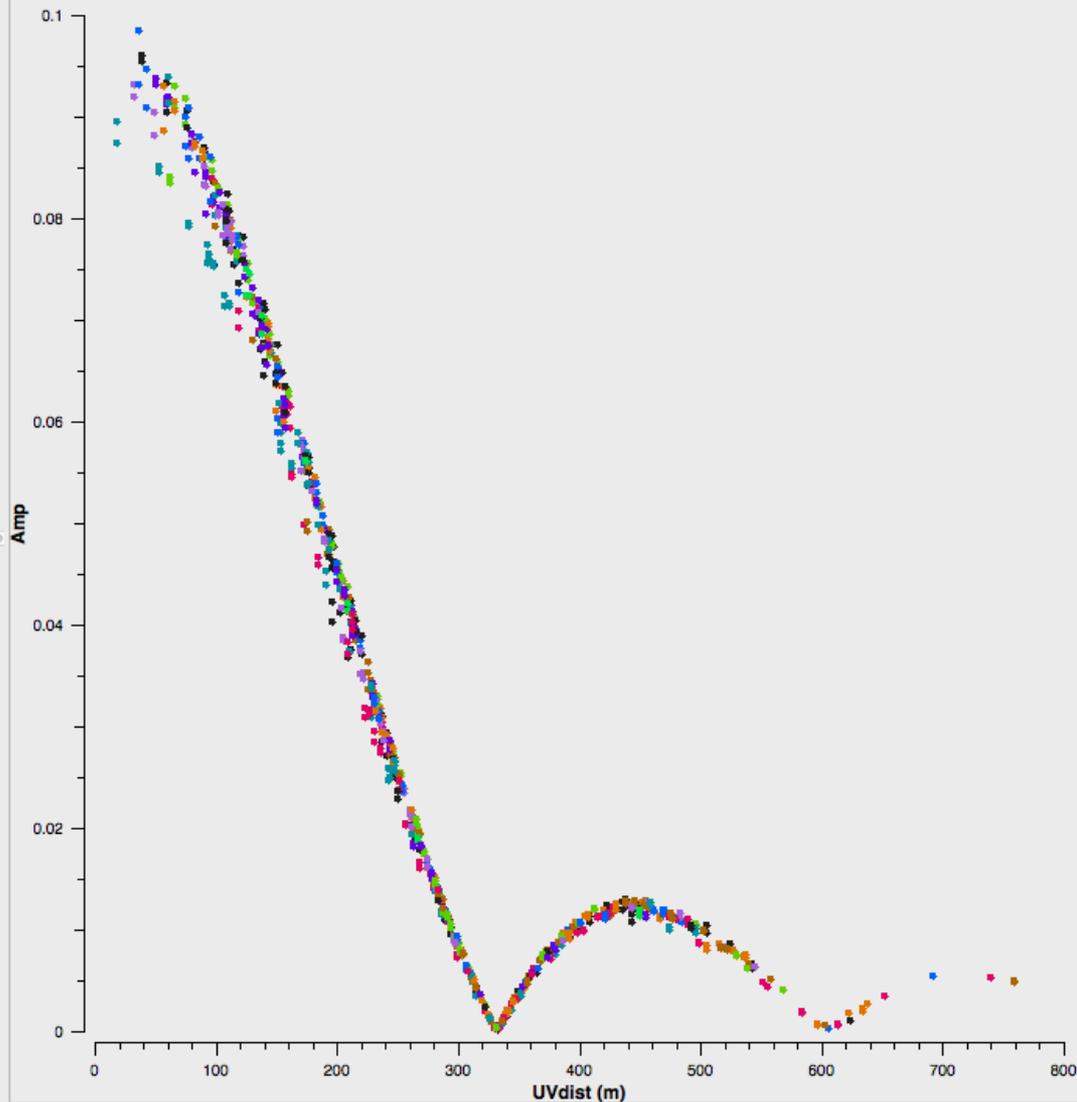
Add Y Axis Data

Add Plot



Reload

Plot



Hold Drawing

Amp vs. UVdist Spw: 0

Iteration

Axis: Spw

Global Axis Scale: X Y

Shared Axis: X Y

Page Header

Contents

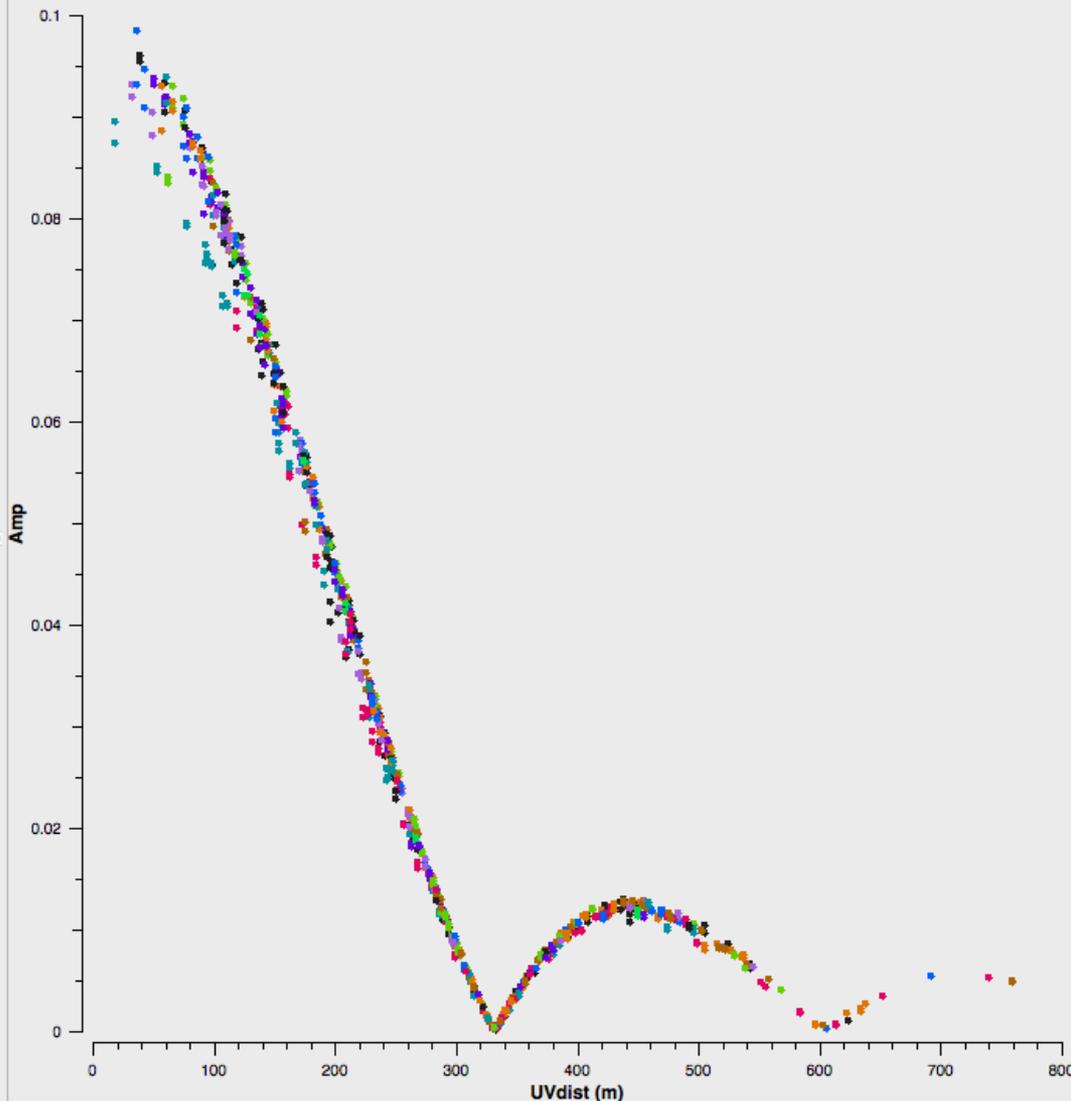


Filename
Y Column(s)
Observation Start Date
Observation Start Time
Observer
Project ID

Canvas Display Transform Page Axes Calibration Data

Add Plot

Reload Plot



Navigation toolbar with icons for zoom, pan, and drawing tools, followed by the text 'Hold Drawing'.

Y Axis Data: Amp

Colorize: Antenna1

Unflagged Points Symbol

None Default

Custom

Style: 2 px, autoscaling

Fill: 0000ff fill

Outline: None Default

Flagged Points Symbol

None Default

Custom

Style: 2 px, circle

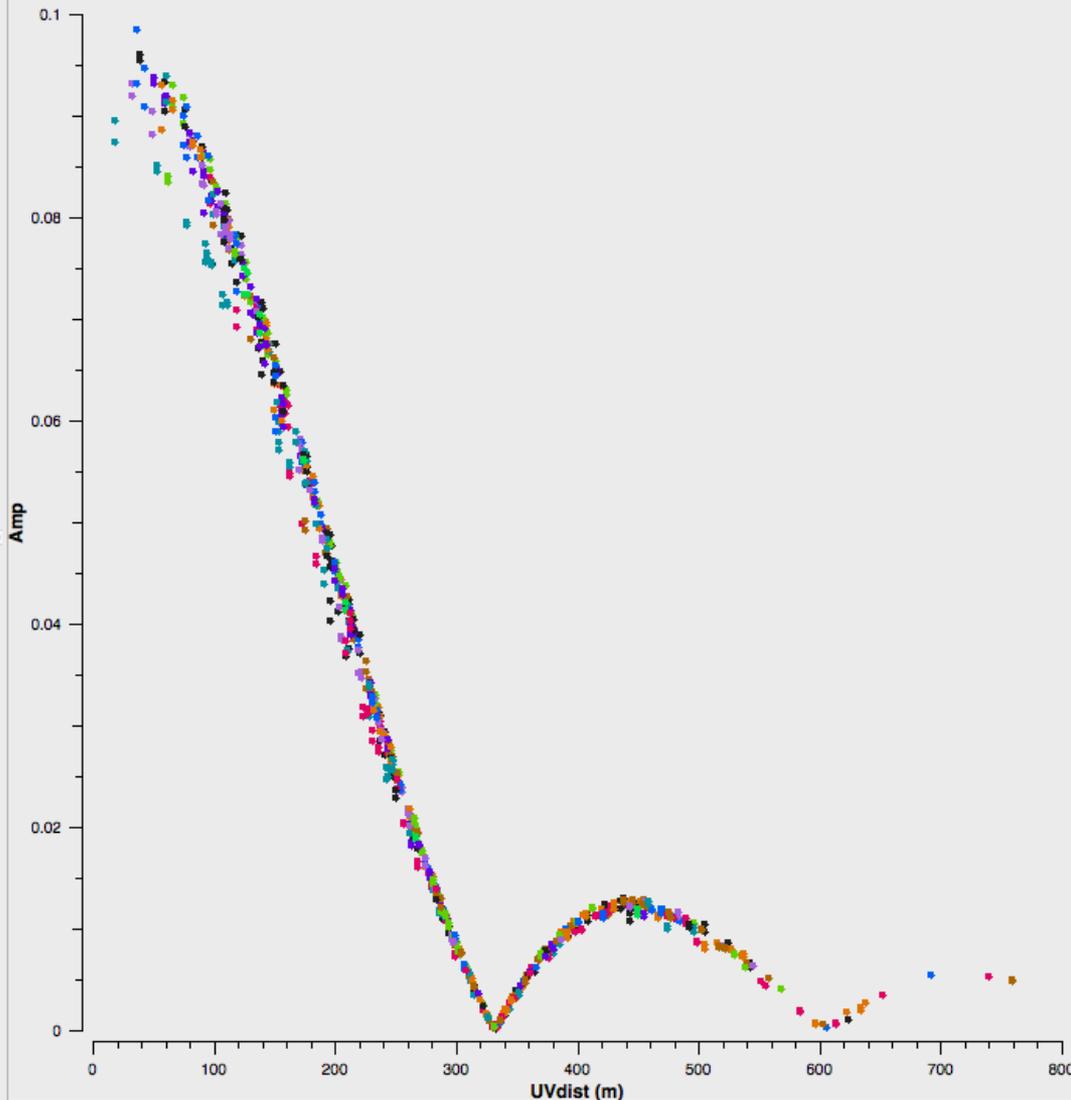
Fill: ff0000 fill

Outline: None Default

Canvas Display Transform Page Axes Calibration Data

Add Plot

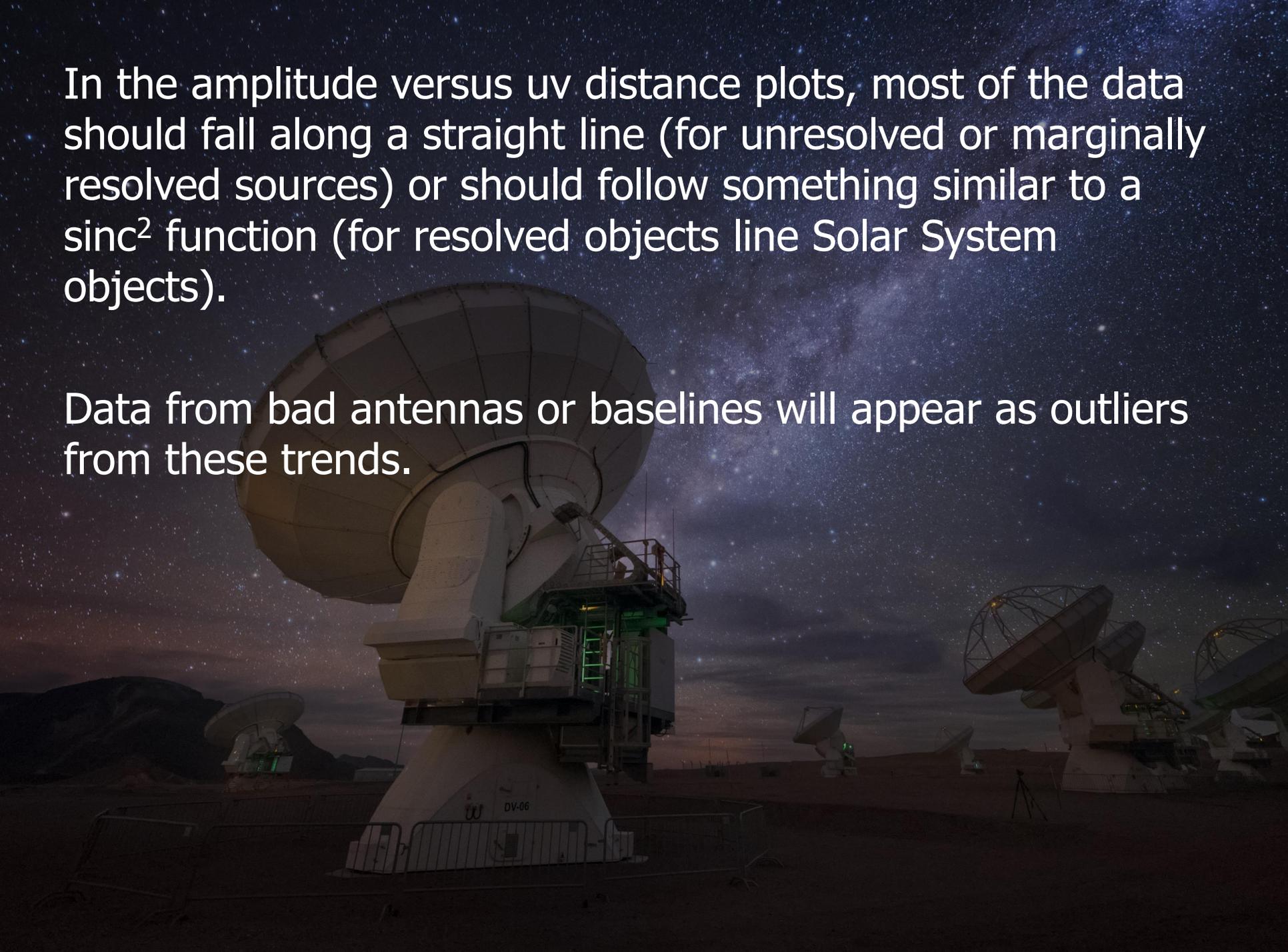
Reload Plot



Navigation and tool icons including zoom, pan, and drawing tools, with a 'Hold Drawing' label.

In the amplitude versus uv distance plots, most of the data should fall along a straight line (for unresolved or marginally resolved sources) or should follow something similar to a sinc^2 function (for resolved objects like Solar System objects).

Data from bad antennas or baselines will appear as outliers from these trends.



File

544a_X25eb.ms.split

Selection

field

spw

timerange

uvrange

antenna

scan

corr

array

observation

intent

feed

msselect

Averaging

Channel channels

Time seconds

Scan Field

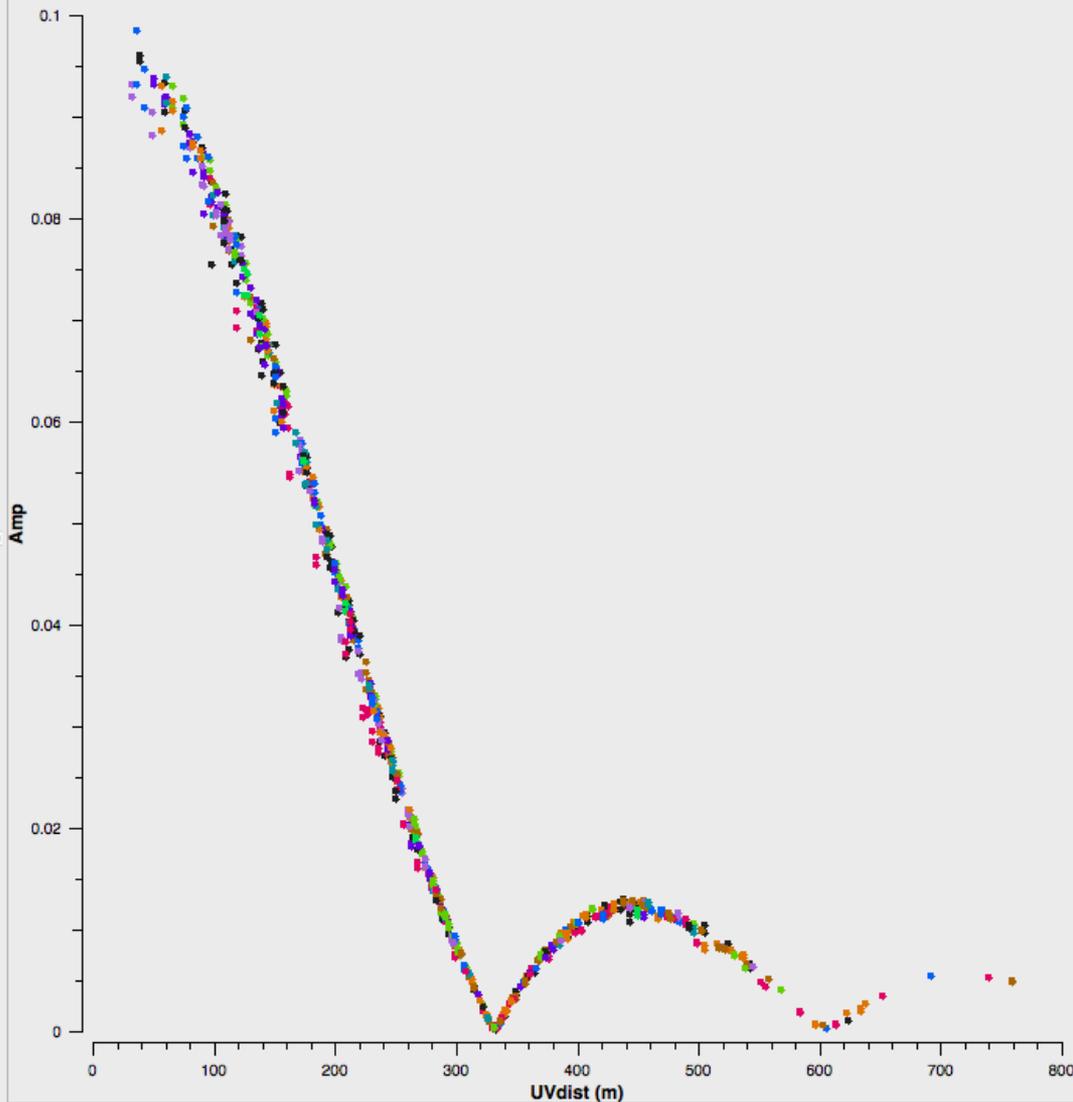
All Baselines Per Antenna

All Spectral Windows

Vector Scalar

Canvas Display Transform Page Axes Calibration Data

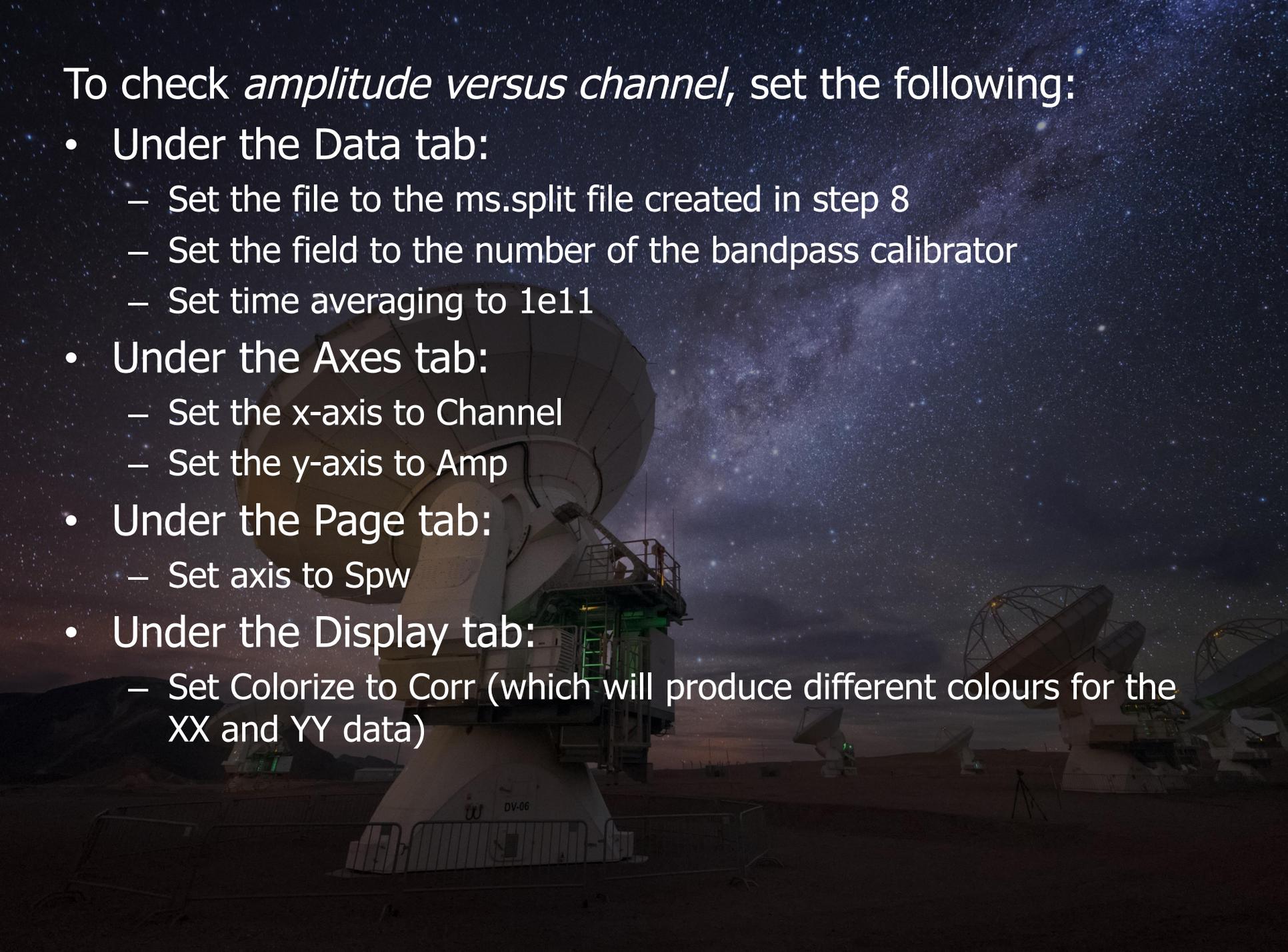
Reload



Navigation toolbar with icons for zoom, pan, home, and other plot controls. Includes a 'Hold Drawing' label.

To check *amplitude versus channel*, set the following:

- Under the Data tab:
 - Set the file to the ms.split file created in step 8
 - Set the field to the number of the bandpass calibrator
 - Set time averaging to 1e11
- Under the Axes tab:
 - Set the x-axis to Channel
 - Set the y-axis to Amp
- Under the Page tab:
 - Set axis to Spw
- Under the Display tab:
 - Set Colorize to Corr (which will produce different colours for the XX and YY data)



Plot Flag Tools Annotate Options

File

544a_X25eb.ms.split

Browse...

Data

Selection

field 0

spw

timerange

uvrange

antenna

scan

corr

array

observation

intent

feed

msselect

Display

Averaging

 Channel channels Time seconds Scan Field All Baselines Per Antenna All Spectral Windows Vector Scalar

Canvas

Transform

Page

Axes

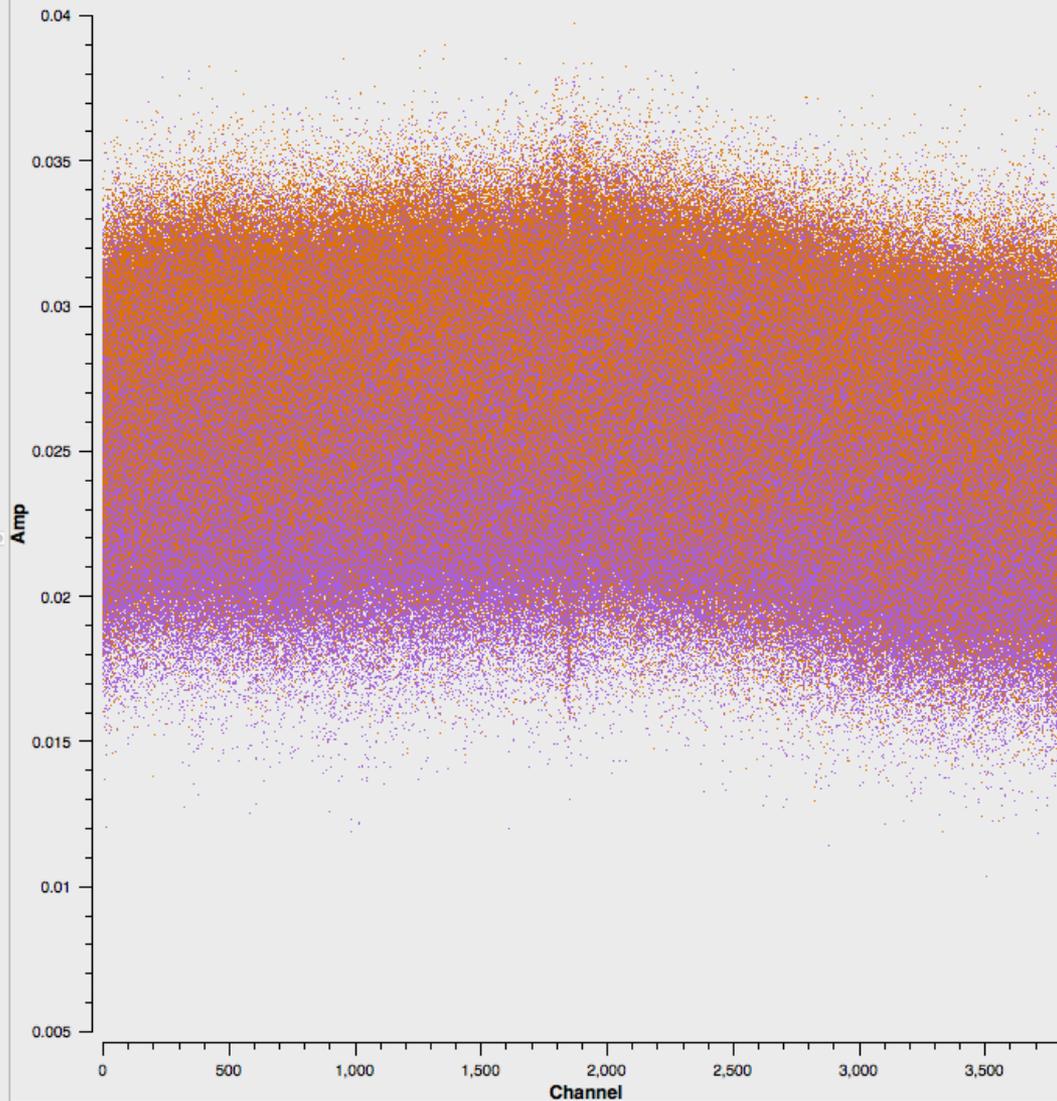
Calibration

Add Plot

 Reload

Plot

Amp vs. Channel Spw: 2



Hold Drawing

X Axis: Channel

Cached: Attach: Bottom Top

Range:

 Automatic

0

to

 0

Data: Amp

Data Column: data

Cached: Attach: Left Right

Range:

 Automatic

0

to

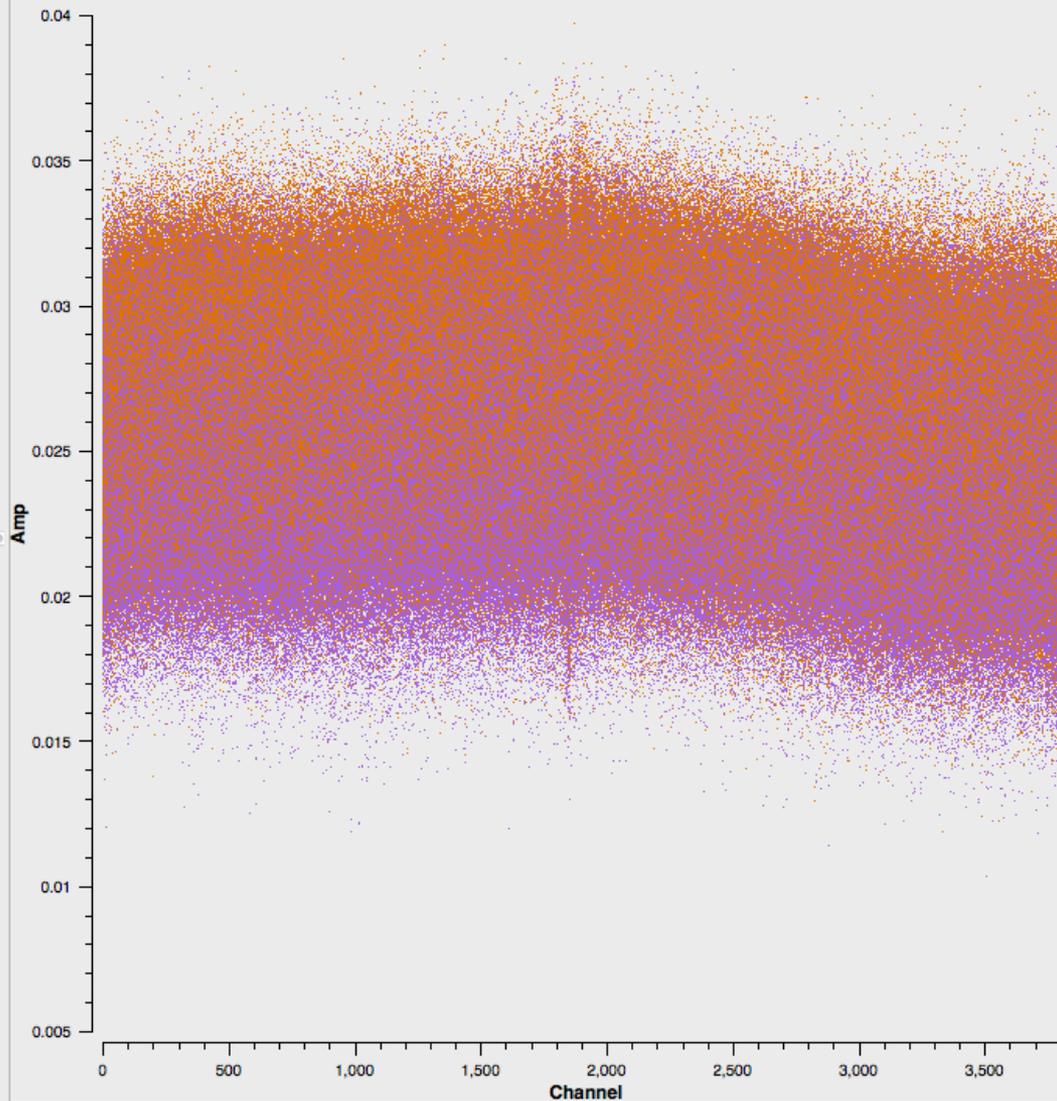
 0

Add Y Axis Data

Add Plot

 Reload

Plot



Hold Drawing

Amp vs. Channel Spw: 2

Iteration

Axis:

Global Axis Scale: X Y

Shared Axis: X Y

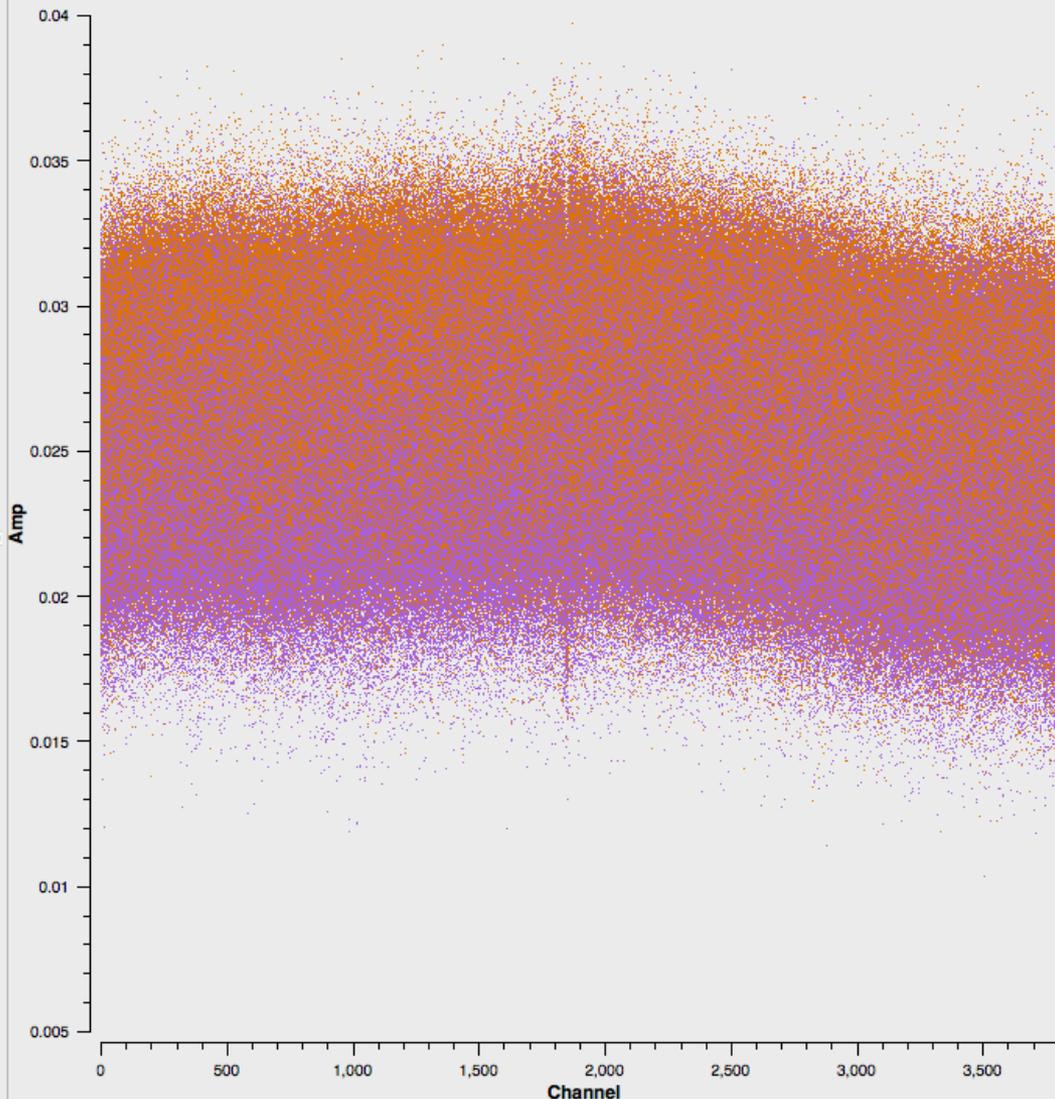
Page Header

Contents



Filename
Y Column(s)
Observation Start Date
Observation Start Time
Observer
Project ID

Canvas Display Transform Page Axes Calibration Data



Add Plot

Reload Plot



Hold Drawing

Plot Flag Tools Annotate Options

Y Axis Data: Amp

 Colorize: Corr

Unflagged Points Symbol

 None Default Custom

Style: 2 px, autoscaling

Fill: 0000ff fill

Outline: None Default

Flagged Points Symbol

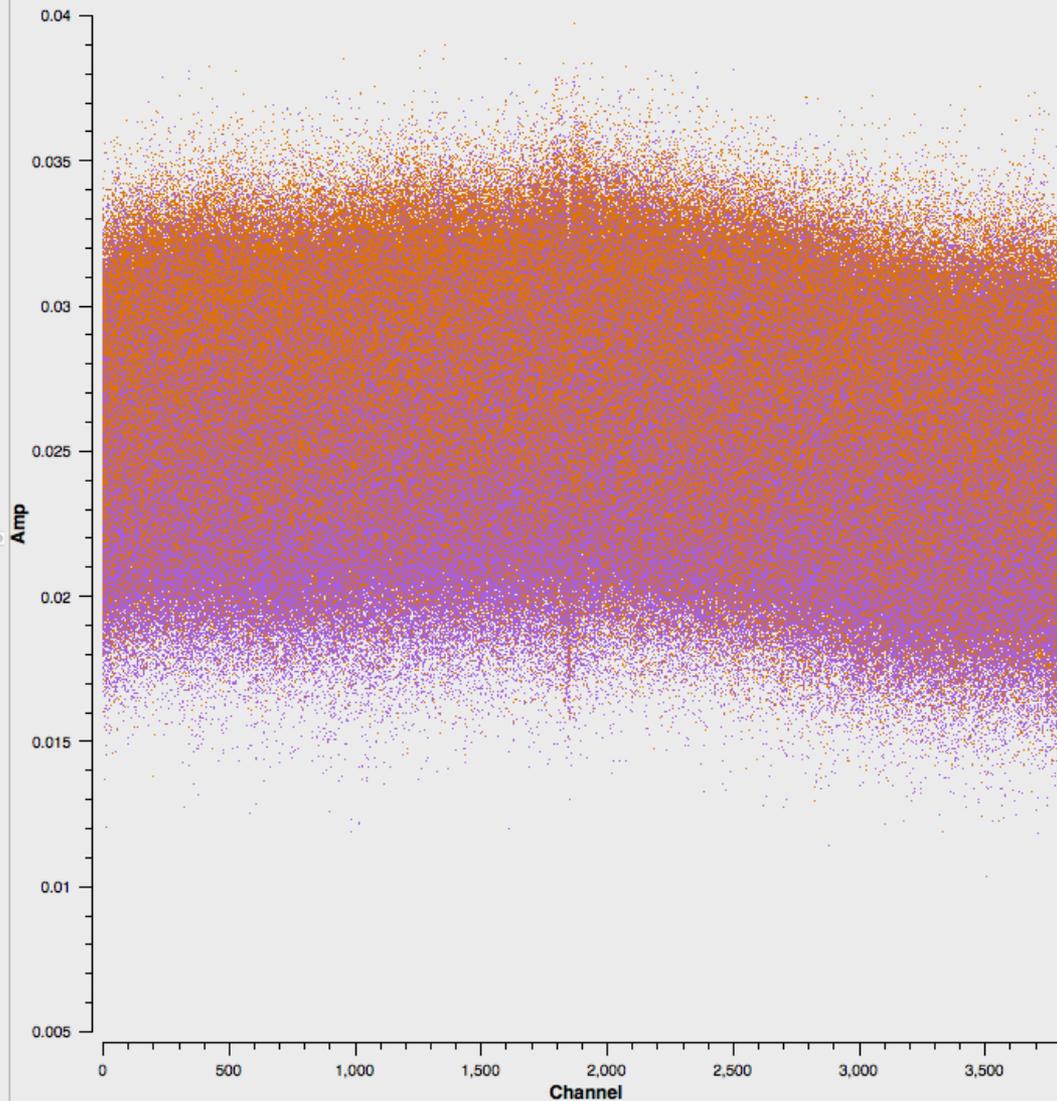
 None Default Custom

Style: 2 px, circle

Fill: ff0000 fill

Outline: None DefaultData
Calibration
Axes
Page
Transform
Display
Canvas

Amp vs. Channel Spw: 2



Add Plot

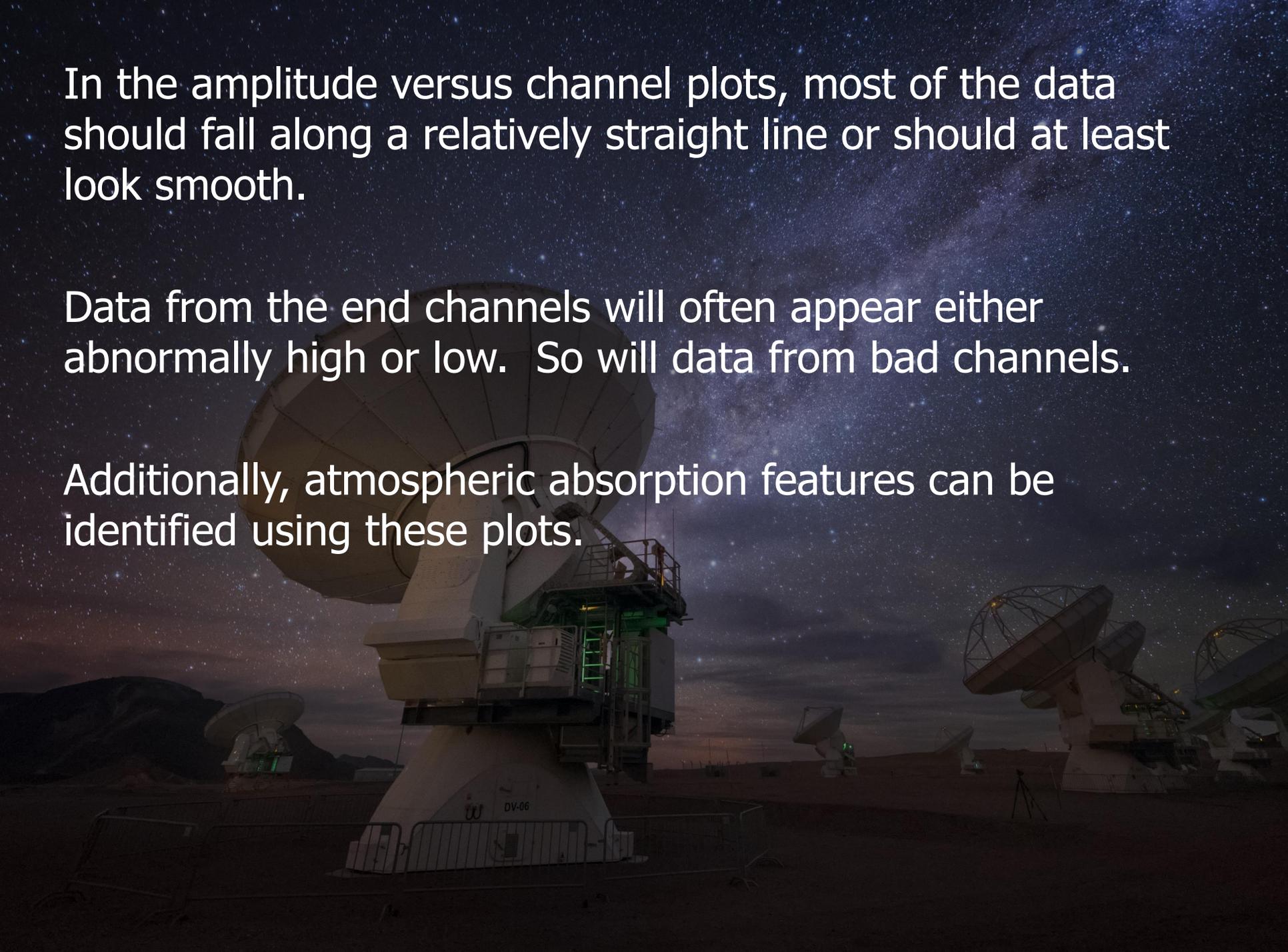
 Reload Plot

Hold Drawing

In the amplitude versus channel plots, most of the data should fall along a relatively straight line or should at least look smooth.

Data from the end channels will often appear either abnormally high or low. So will data from bad channels.

Additionally, atmospheric absorption features can be identified using these plots.



Plot Flag Tools Annotate Options

File

544a_X25eb.ms.split

Browse...

Data

Selection

field 0

spw

timerange

uvrange

antenna

scan

corr

array

observation

intent

feed

msselect

Page

Transform

Display

Canvas

Averaging

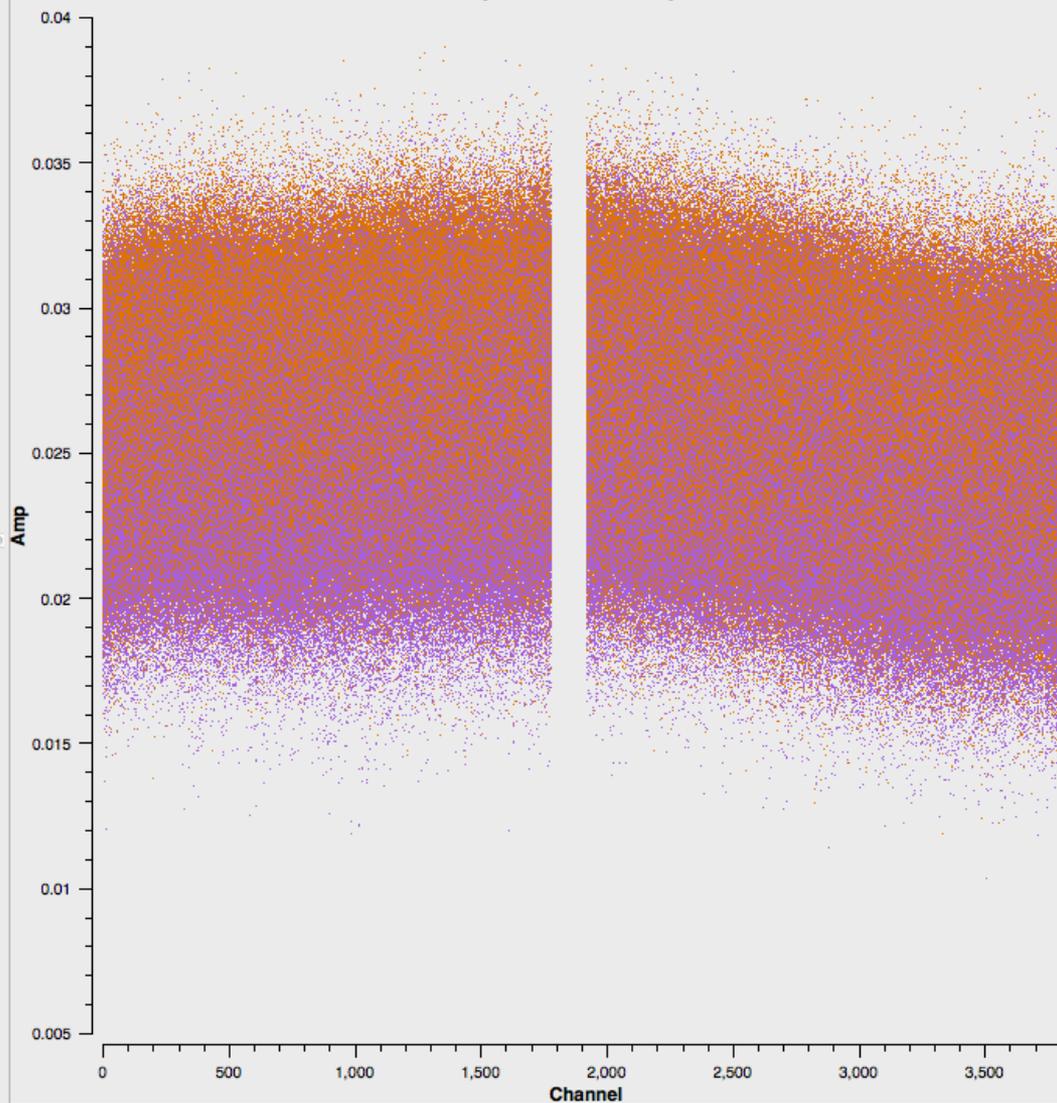
 Channel channels Time seconds Scan Field All Baselines Per Antenna All Spectral Windows Vector Scalar

Add Plot

 Reload

Plot

Amp vs. Channel Spw: 2



Hold Drawing

To check *phase versus scan*, set the following:

- Under the Data tab:
 - Set the file to the ms.split file created in step 8
 - Set the field to the number of the phase calibrator
 - Set the spw to one of the spw numbers (each one should be checked separately)
 - Set the antenna to the reference antenna followed by "&*"
 - Set channel averaging to 4100
- Under the Axes tab:
 - Set the x-axis to Scan
 - Set the y-axis to Phase and set the range to 180 to -180
- Under the Page tab:
 - Set axis to Baseline
- Under the Display tab:
 - Set Colorize to Corr (which will produce different colours for the XX and YY data)

File

544a_X25eb.ms.split

Browse...

Data

Selection

field 2

spw 0

timerange

uvrange

antenna DV11&*

scan

corr

array

observation

intent

feed

msselect

Averaging

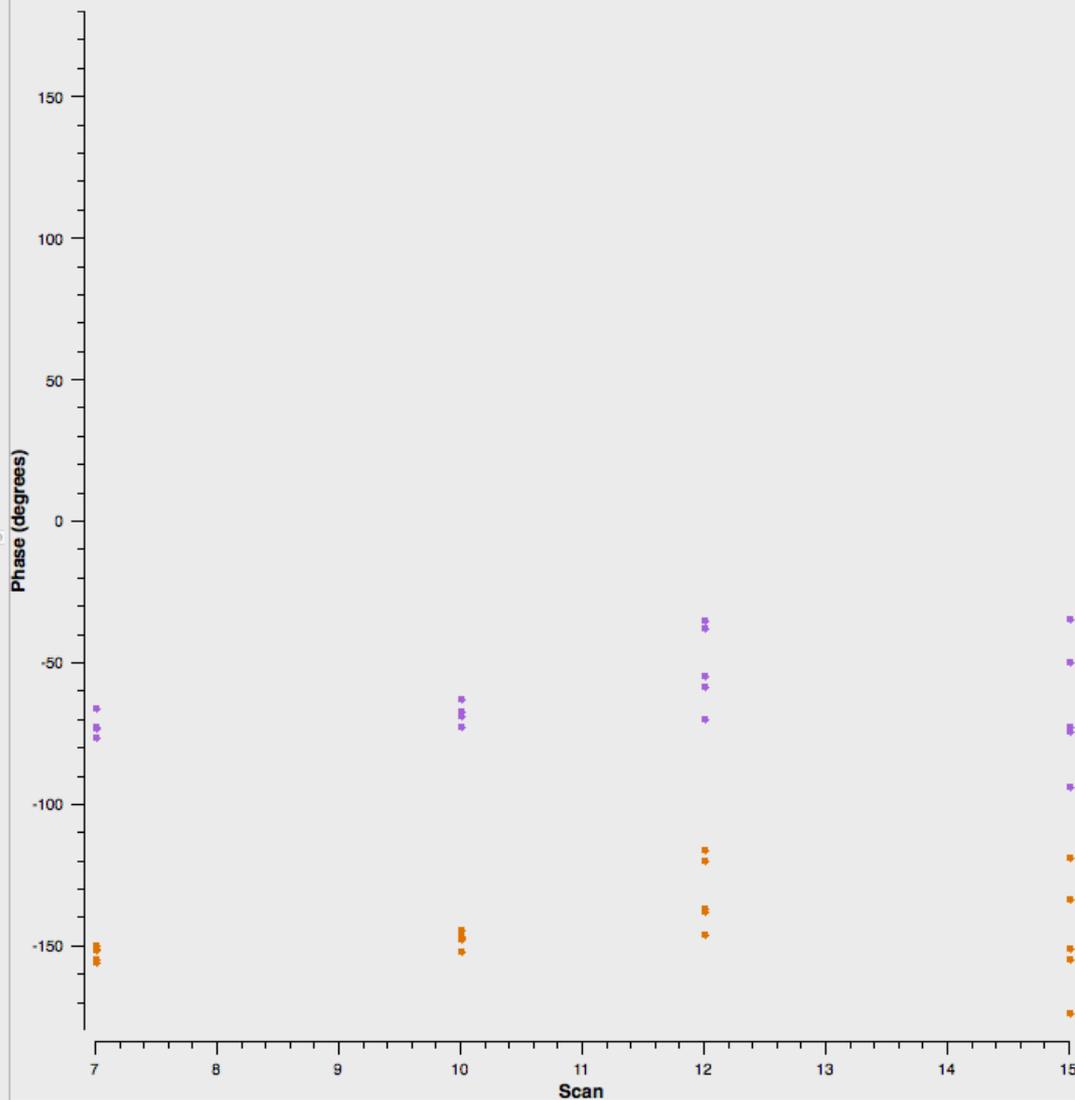
 Channel 4100 channels Time 0 seconds Scan Field All Baselines Per Antenna All Spectral Windows Vector Scalar

Canvas

Add Plot

 Reload

Plot



Hold Drawing

X Axis: Scan

Cached: Attach: Bottom Top

Range:

 Automatic

0

to

 0

Data: Phase

Data Column: data

Cached: Attach: Left Right

Range:

 Automatic

180

to

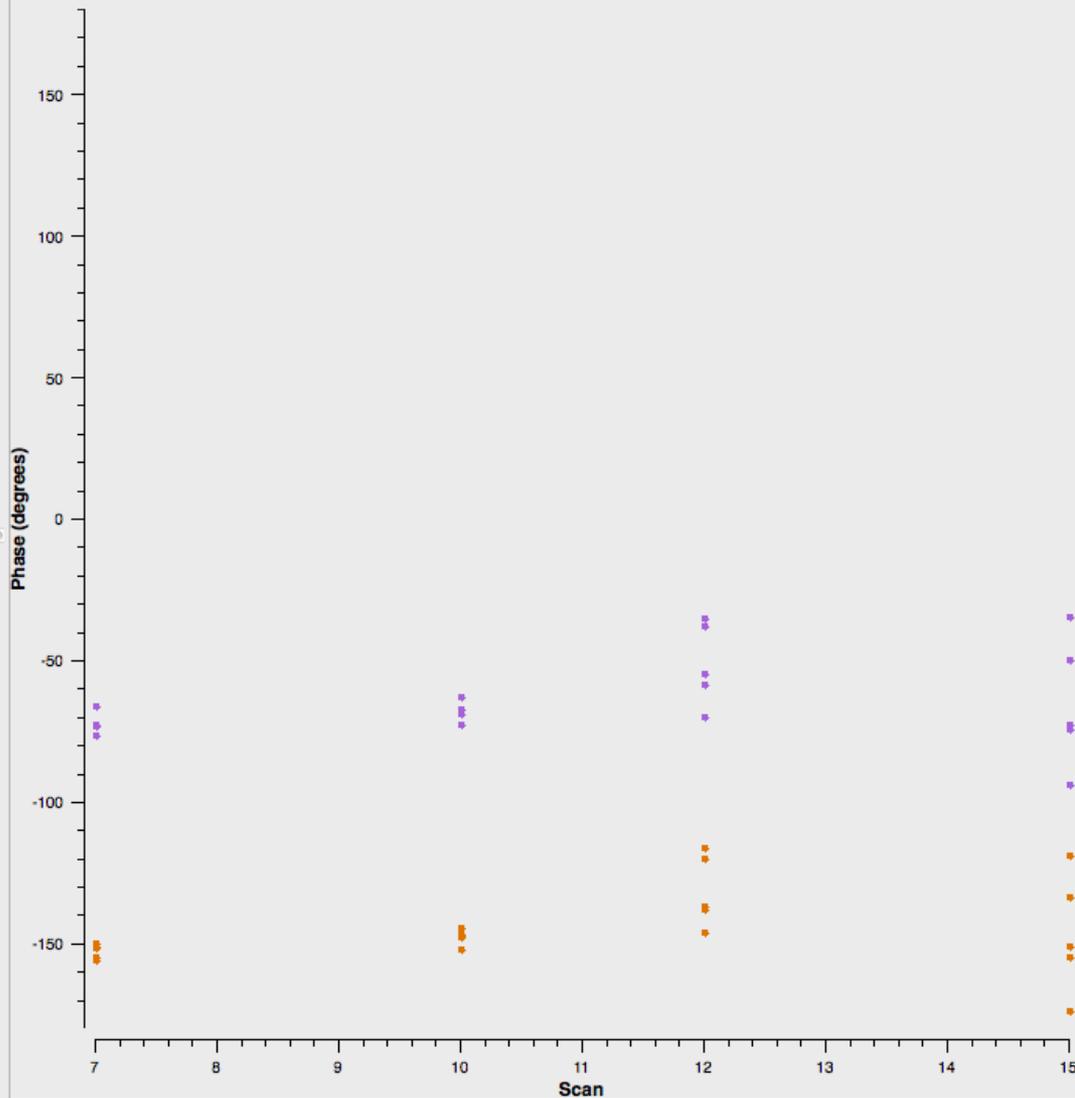
 -180

Add Y Axis Data

Add Plot

Reload

Plot



Hold Drawing

Iteration

Axis: Baseline

Global Axis Scale: X Y

Shared Axis: X Y

Page Header

Contents

Filename

Y Column(s)

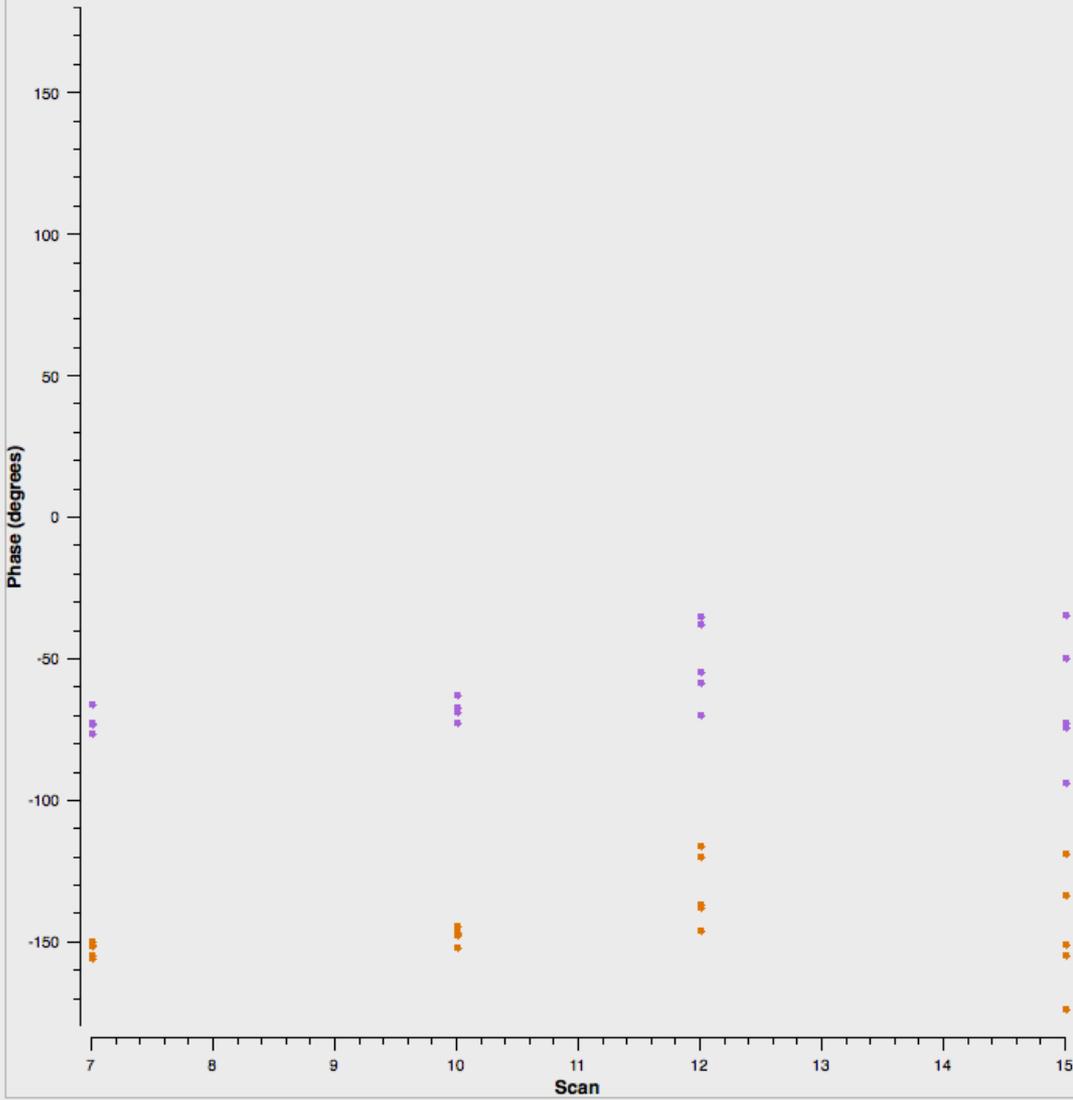
Observation Start Date

Observation Start Time

Observer

Project ID

Canvas Display Transform Page Axes Calibration Data



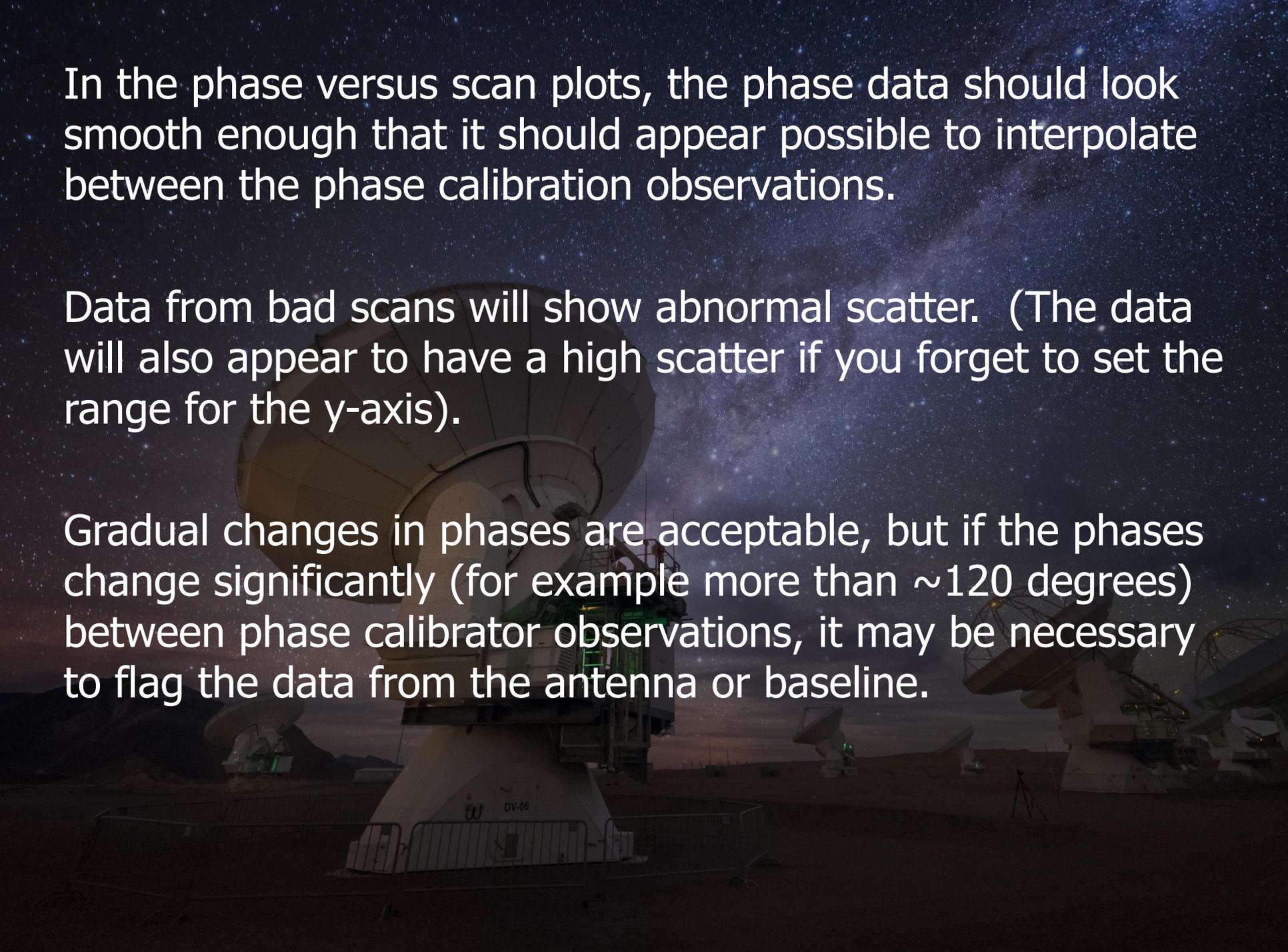
Add Plot Reload Plot

Navigation icons: Home, Back, Forward, Refresh, Zoom, Pan, and other plot controls.

In the phase versus scan plots, the phase data should look smooth enough that it should appear possible to interpolate between the phase calibration observations.

Data from bad scans will show abnormal scatter. (The data will also appear to have a high scatter if you forget to set the range for the y-axis).

Gradual changes in phases are acceptable, but if the phases change significantly (for example more than ~ 120 degrees) between phase calibrator observations, it may be necessary to flag the data from the antenna or baseline.



Step 11: Put a model for the flux calibrator into the measurement set

This step adds information indicating the flux density of one of the sources in the data (typically the amplitude calibrator, although a different source can be used if necessary).

Either a Solar System object or a quasar is used in this step.

The flux densities for Solar System objects are based on well-calibrated models (although the Ceres and Pallas models are suspect).

The quasar flux densities are based on a catalog of periodic measurements calibrated against Solar System objects.

Step 12: Save flags

This is one of several steps where the flagging information is saved.

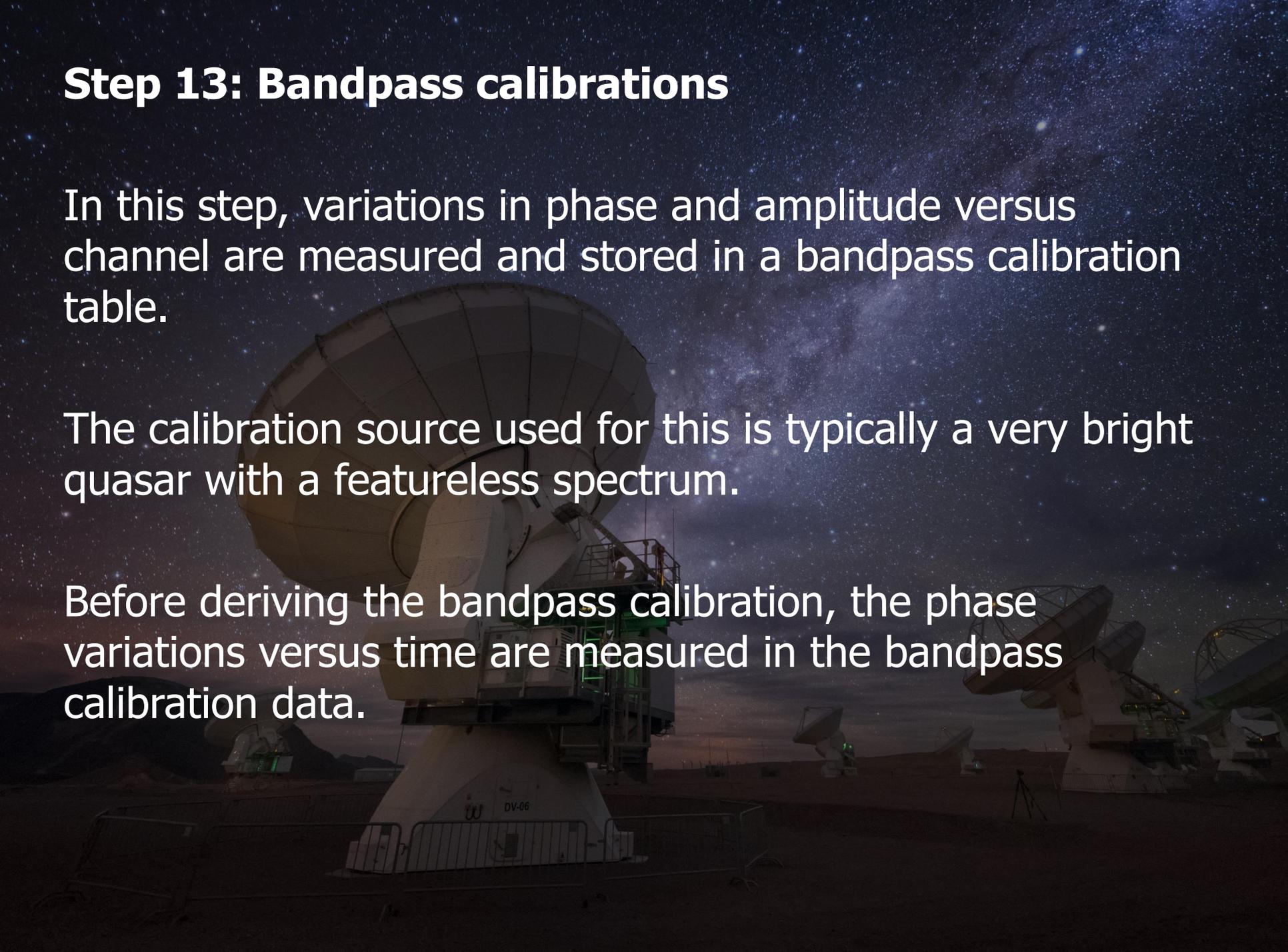


Step 13: Bandpass calibrations

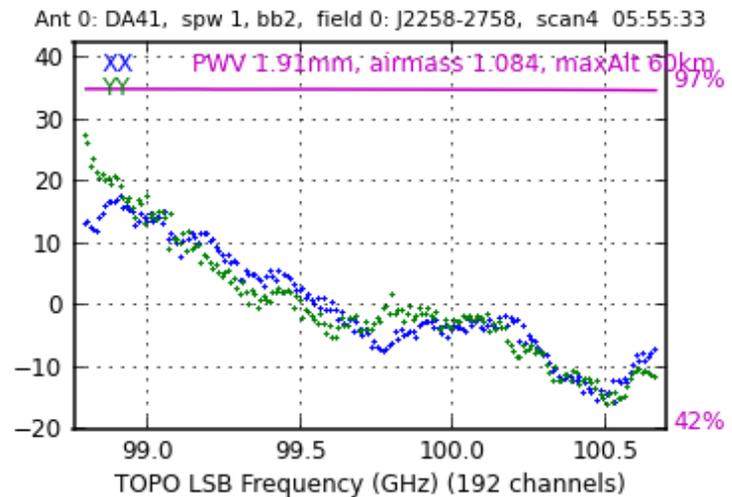
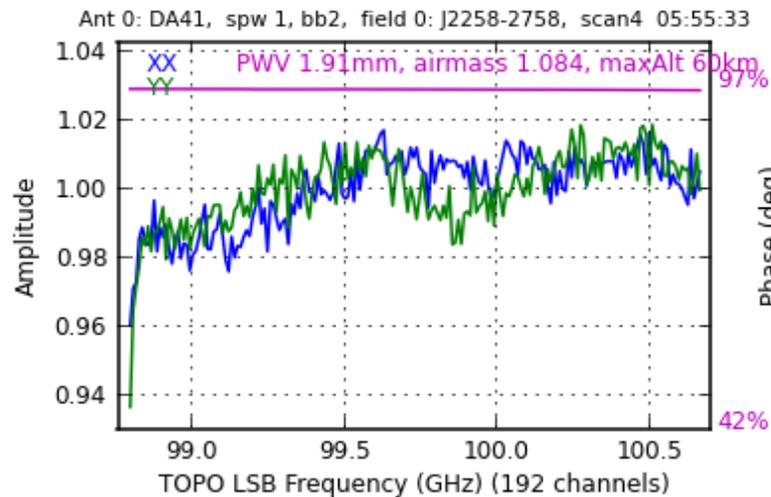
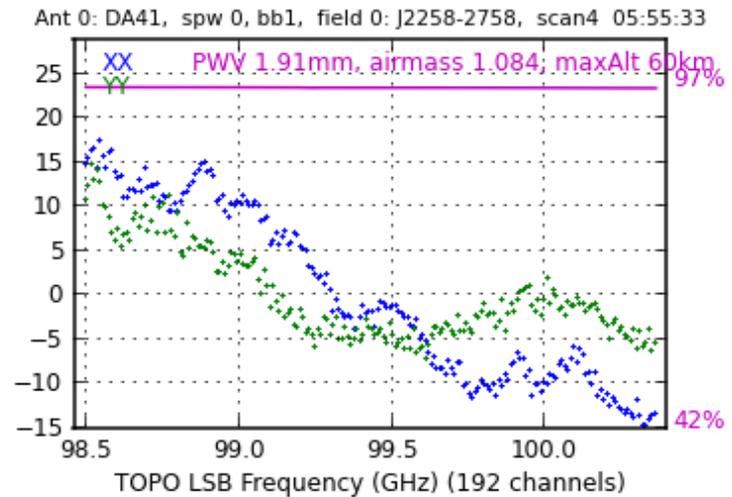
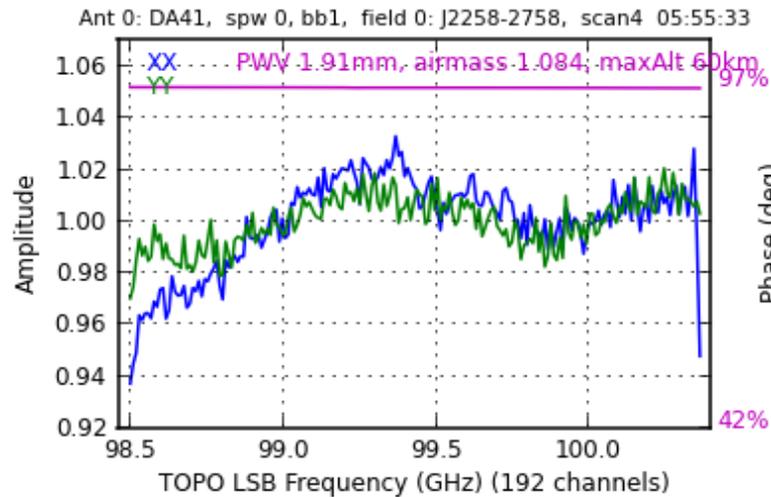
In this step, variations in phase and amplitude versus channel are measured and stored in a bandpass calibration table.

The calibration source used for this is typically a very bright quasar with a featureless spectrum.

Before deriving the bandpass calibration, the phase variations versus time are measured in the bandpass calibration data.



uid__A002_X87544a_X25eb.ms.split.bandpass_smooth20ch



Step 14: Save flags

This is one of several steps where the flagging information is saved.



Step 15: Gain calibrations

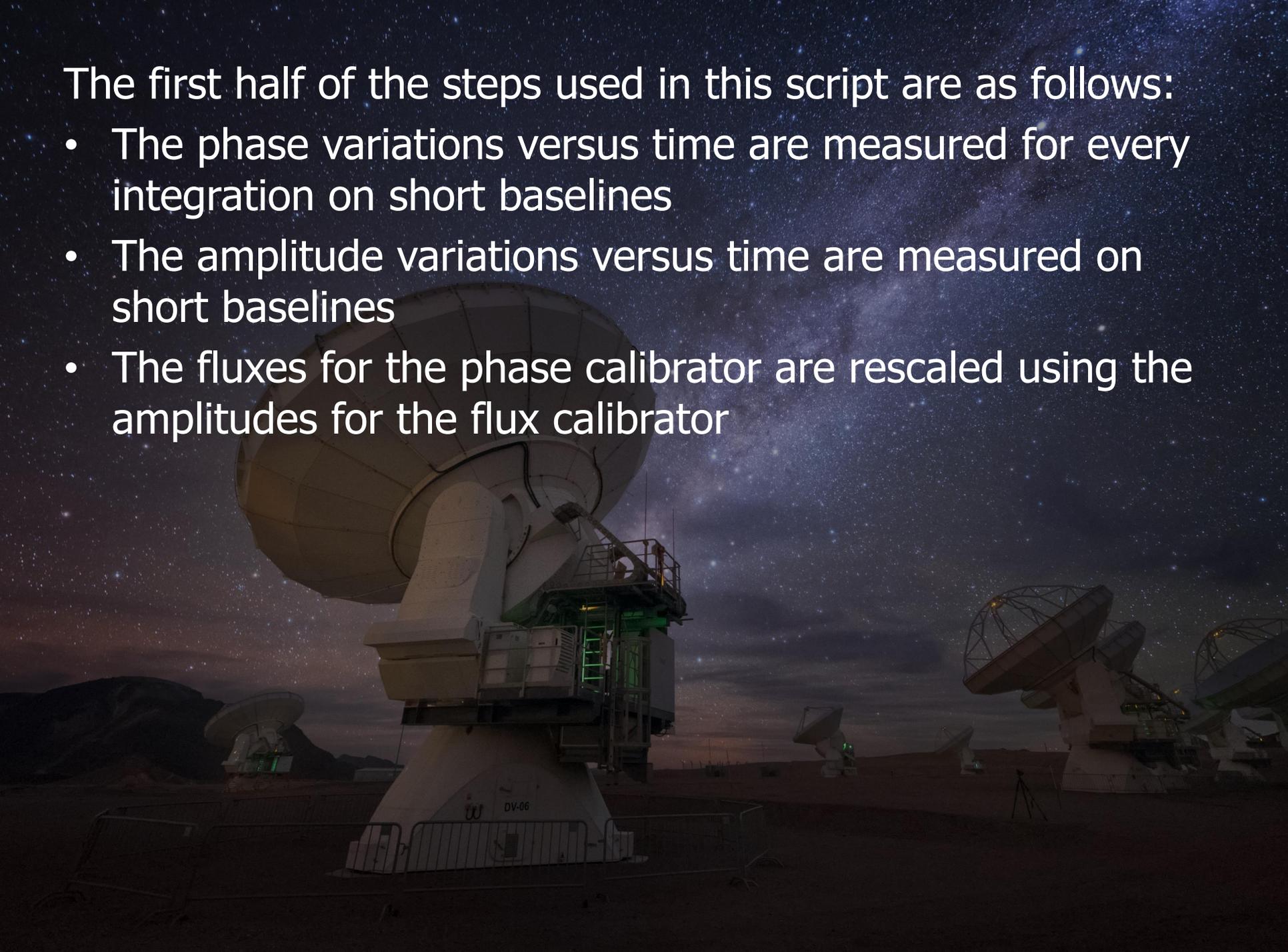
In this step, variations in phase and amplitude versus time are measured and stored in multiple calibration tables.

This step is complicated by the use of a resolved object as a flux calibrator. The Solar System object has more reliably-modelled fluxes, but the signal is not detected on some of the longer baselines.

To properly calibrate the data, the flux calibration for the phase calibrator needs to be derived using the short baselines data for the flux calibrator. After that, the phase and amplitude calibration versus time can be derived for the science fields.

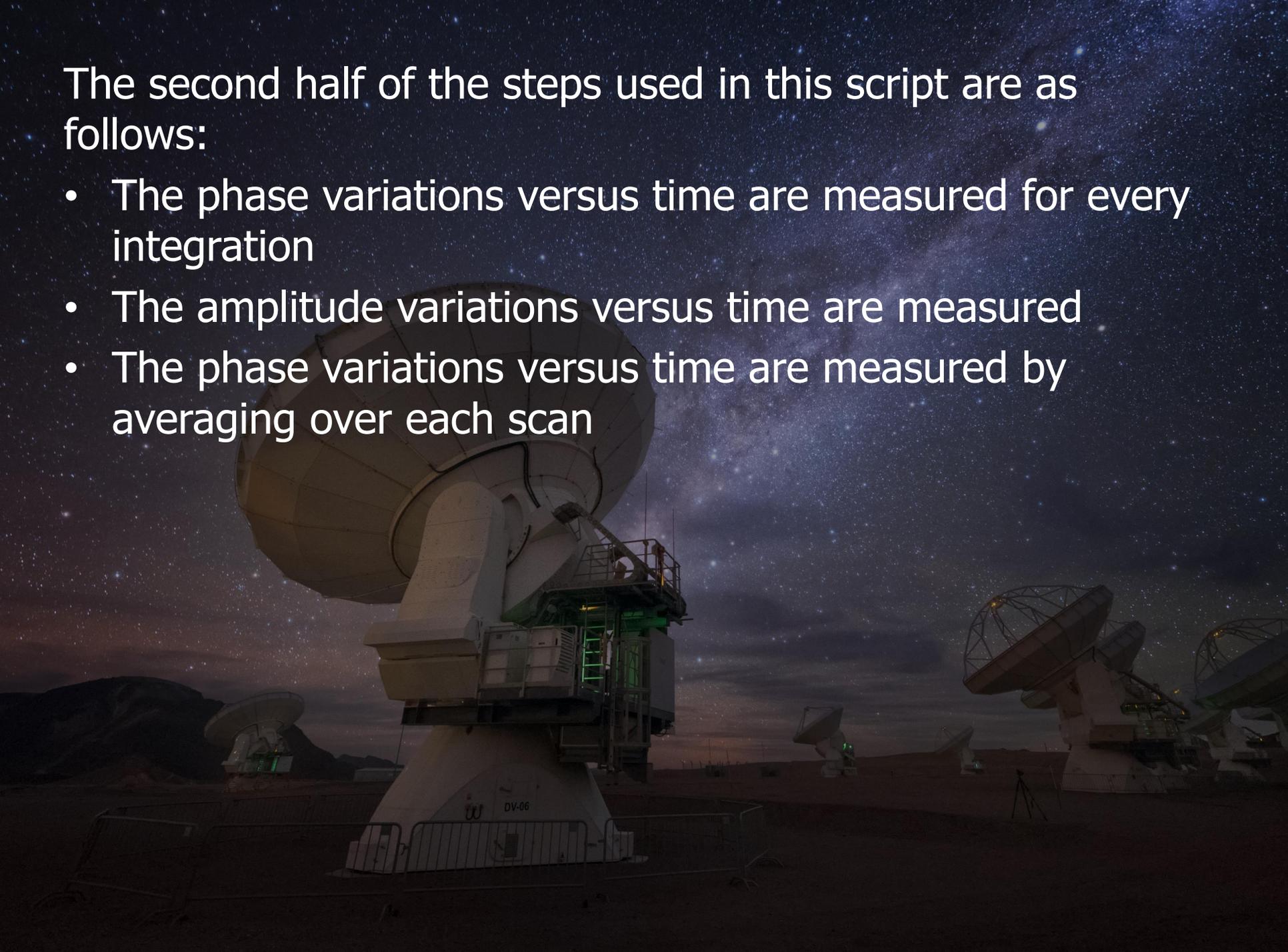
The first half of the steps used in this script are as follows:

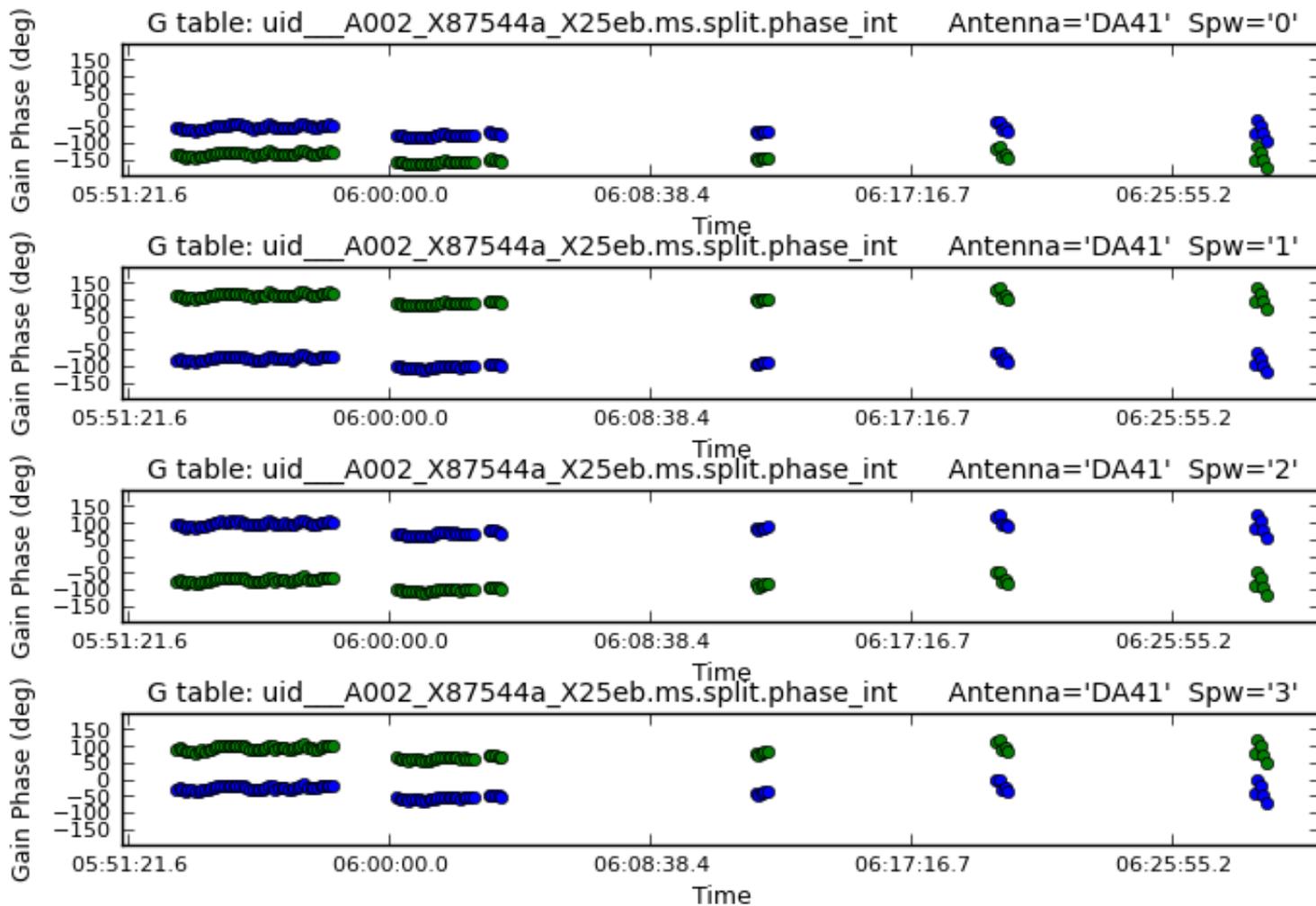
- The phase variations versus time are measured for every integration on short baselines
- The amplitude variations versus time are measured on short baselines
- The fluxes for the phase calibrator are rescaled using the amplitudes for the flux calibrator

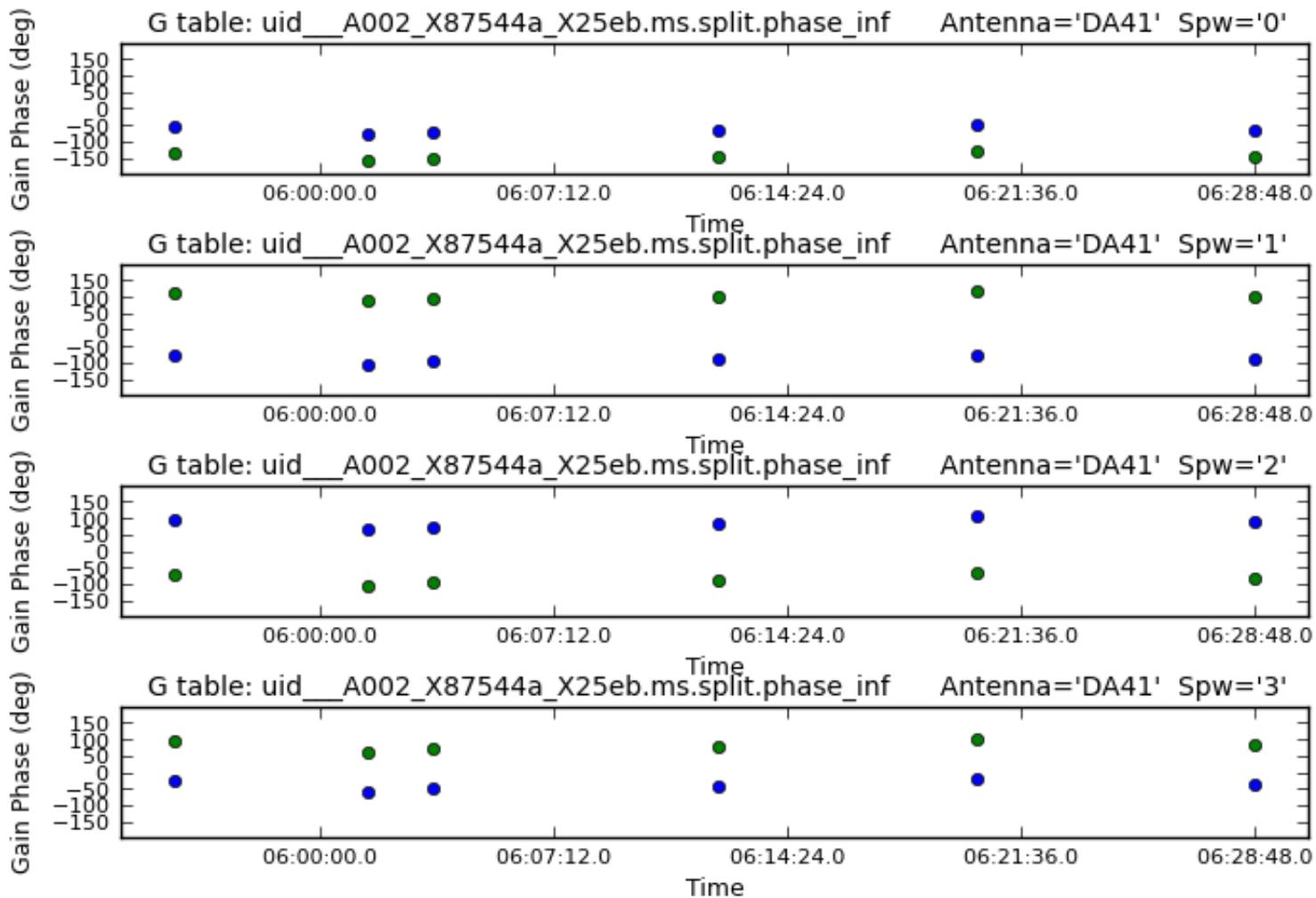


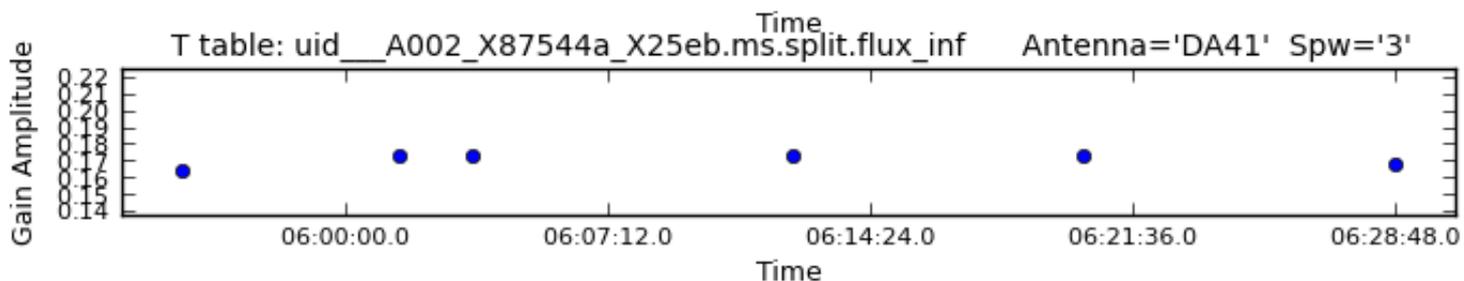
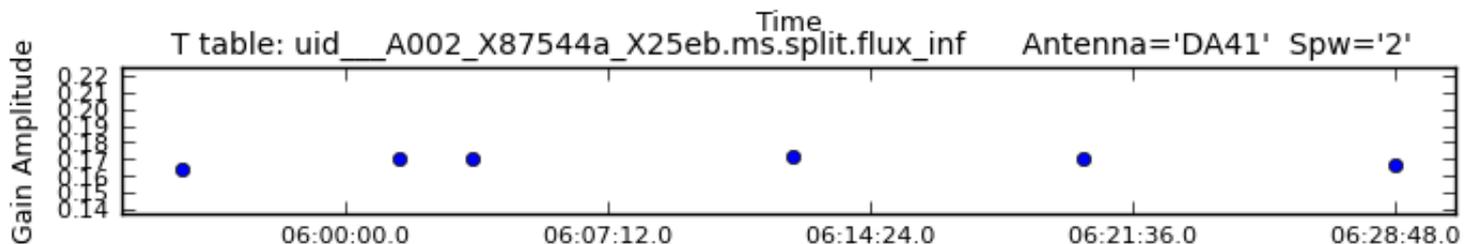
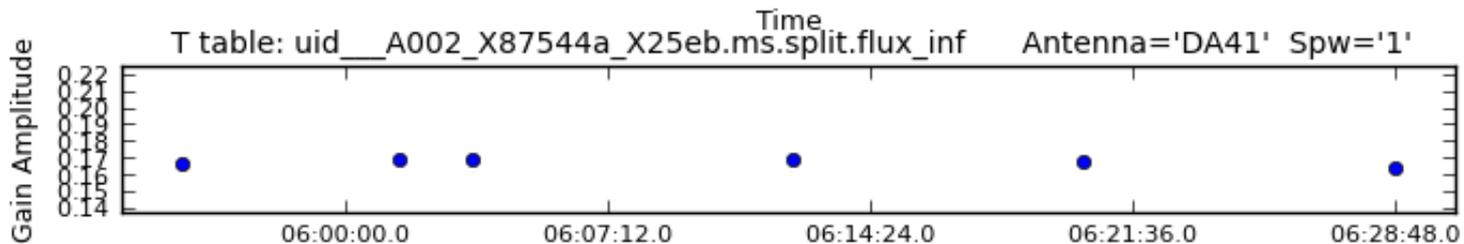
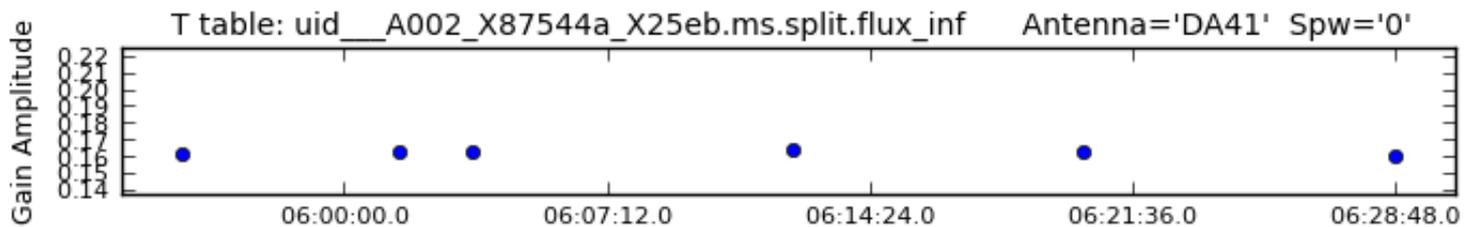
The second half of the steps used in this script are as follows:

- The phase variations versus time are measured for every integration
- The amplitude variations versus time are measured
- The phase variations versus time are measured by averaging over each scan









Step 16: Save flags

This is one of several steps where the flagging information is saved.



Step 17: Application of the bandpass and gain calibration tables

The tables generated in steps 13 and 15 are applied to the data.



Step 18: Split out the corrected column

The fully corrected data are separated from the rest of the data and put into a new measurement set.



Step 19: Save flags

This is one of several steps where the flagging information is saved.

