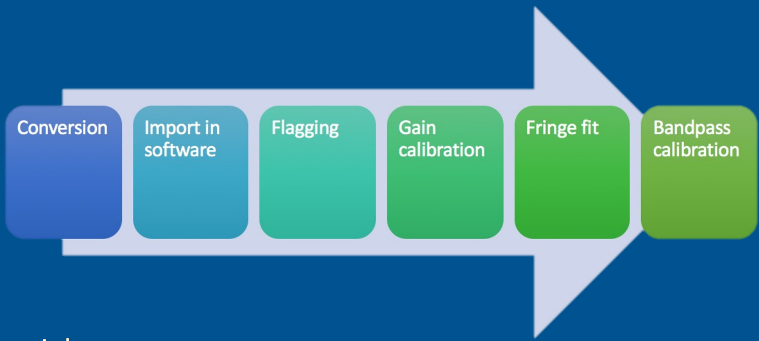
VLBI continuum calibration

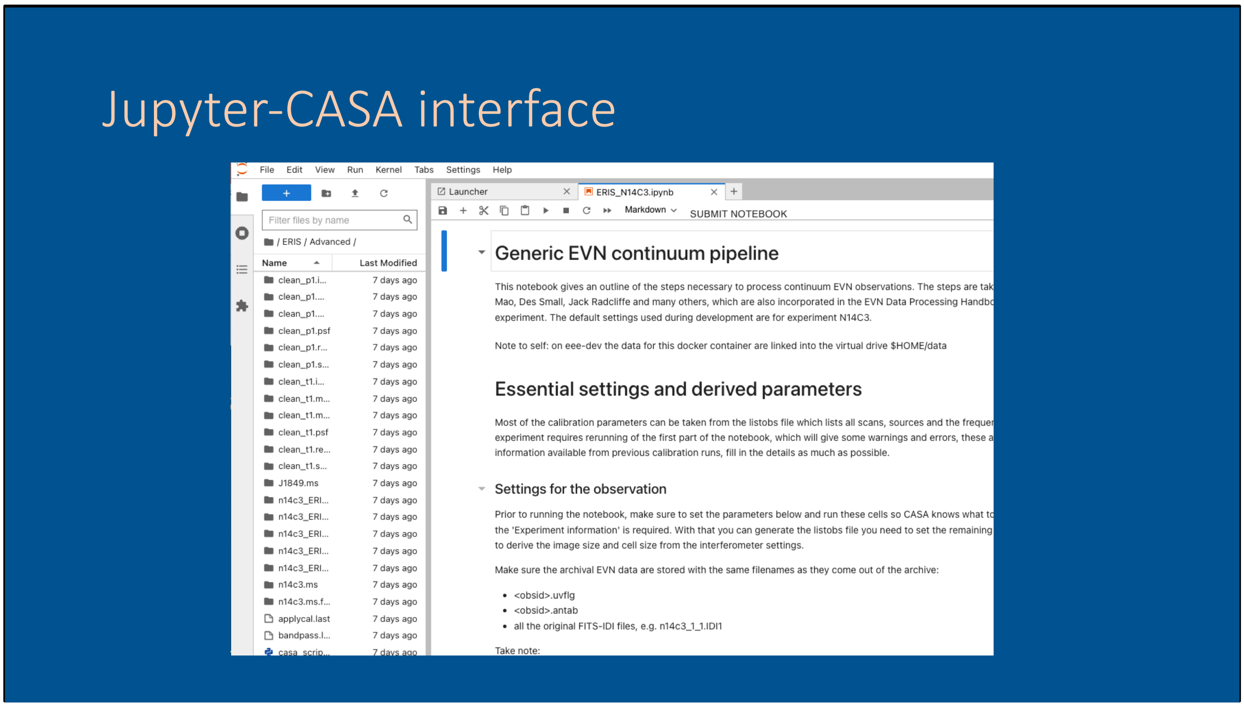
Getting your hands dirty

All calibration steps (tomorrow!)

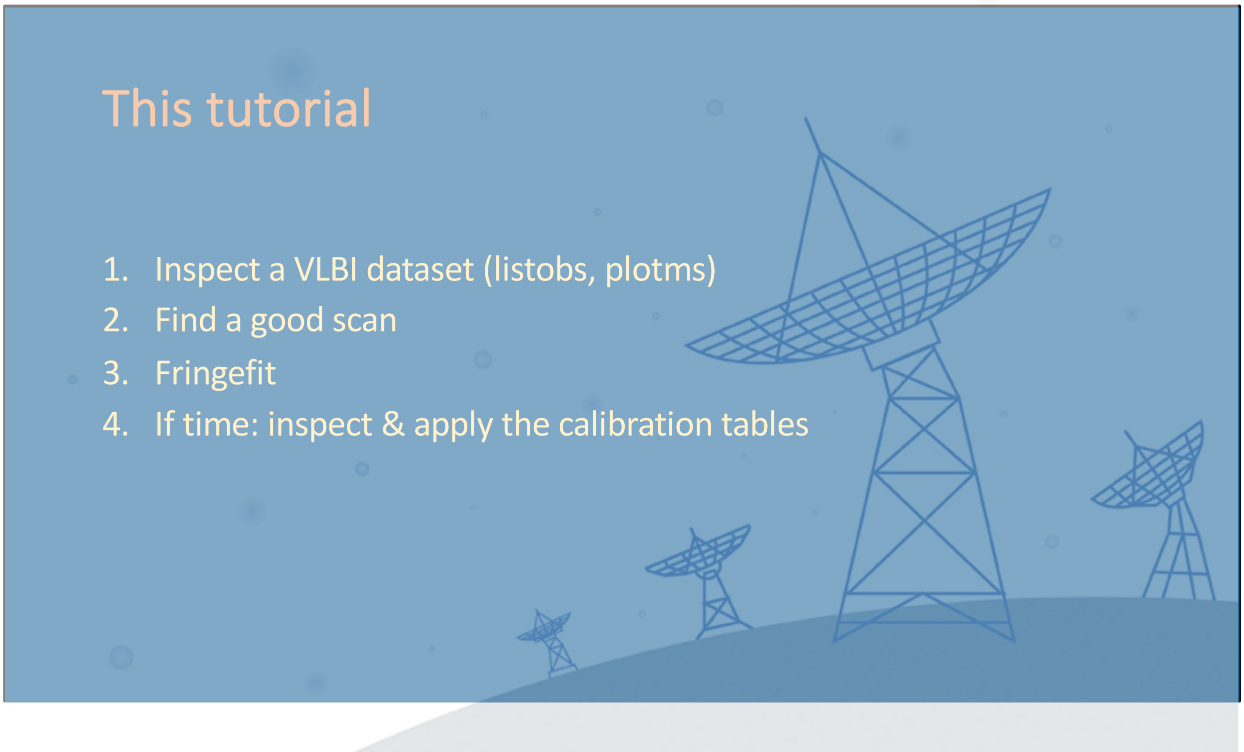


Online materials:
Gitea repository
DARA EVN continuum tutorial

This tutorial is a summary. It is strongly recommended to use the original DARA EVN tutorial in the link for a full VLBI calibration tutorial.



Within JIVE we have developed a Jupyter-CASA interface, allowing users to calibrate their data in a Jupyter environment.



In this plenary tutorial we will focus on the fringe fitting step. A precalibrated dataset has been provided, as well as the Tsys and gain-curve calibration tables required.

This tutorial

Preparations:

- Have the materials downloaded from Gitea
- Make sure you type `git pull` before starting
- Go into the Plenary folder
- Launch CASA

You can execute every command in a WHITE BOX
(beware of plotms!)

1. Inspecting your data

CASA task `listobs`

```
# listobs -- List the summary of a data set in the logger or i
vis             = ''                        # Name of input visib
selectdata      = True                      # Data selection para
    spw         = ''                        # Selection based on
    field       = ''                        # Selection based on
    antenna     = ''                        # Selection based on
    uvrange     = ''                        # Selection based on
    timerange   = ''                        # Selection based on
    correlation = ''                        # Selection based on
    scan        = ''                        # Selection based on
    intent      = ''                        # Selection based on
    feed        = ''                        # Selection based on
    array       = ''                        # Selection based on
    observation = ''                        # Selection based on
verbose         = True                      # Controls level of i
listfile        = ''                        # Name of disk file t
listunfl        = False                     # List unflagged row
cachesize       = 50.0                      # EXPERIMENTAL. Maxim
```

These are the default paramteres for listobs. It is recommended to write the output to file to ensure you have this for future reference.

```
listobs(vis='n14c3_TY_GC_flags.ms/')
```

```
================================================================================
           MeasurementSet Name:  /Users/bemmel/Science/EVN_CASA/ERIS/Plenary/n14c3_TY_GC_flags.ms      MS Version 2
================================================================================
   Observer: N14C3     Project: N14C3
Observation: EVN
Data records: 1846080       Total elapsed time = 10800 seconds
   Observed from   22-Oct-2014/12:00:00.0   to   22-Oct-2014/15:00:00.0 (UTC)

   ObservationID = 0         ArrayID = 0
  Date        Timerange (UTC)          Scan  FldId FieldName             nRows     SpwIds   Average Interval(s)    ScanIntent
  22-Oct-2014/12:00:00.0 - 12:04:00.0     1      1 3C345                 52800  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:06:00.0 - 12:10:00.0     2      1 3C345                 63360  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:12:00.0 - 12:13:00.0     3      1 3C345                 15840  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:13:40.0 - 12:14:40.0     4      1 3C345                 15840  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:15:20.0 - 12:16:20.0     5      0 J1640+3946            15840  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:17:00.0 - 12:18:00.0     6      1 3C345                 13200  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:18:40.0 - 12:19:40.0     7      0 J1640+3946            13200  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:20:20.0 - 12:21:20.0     8      1 3C345                 15840  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:22:00.0 - 12:23:00.0     9      0 J1640+3946            15840  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:23:40.0 - 12:24:40.0    10      1 3C345                 13200  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:25:20.0 - 12:26:20.0    11      0 J1640+3946            13200  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:27:00.0 - 12:28:00.0    12      1 3C345                 15840  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:28:40.0 - 12:29:40.0    13      0 J1640+3946            15840  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:30:20.0 - 12:31:20.0    14      1 3C345                 13200  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:32:00.0 - 12:33:00.0    15      0 J1640+3946            13200  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:33:40.0 - 12:34:40.0    16      1 3C345                 15840  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:35:20.0 - 12:36:20.0    17      0 J1640+3946            15840  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
              12:37:00.0 - 12:38:00.0    18      1 3C345                 13200  [0,1,2,3,4,5,6,7]  [2, 2, 2, 2, 2, 2, 2, 2]
```

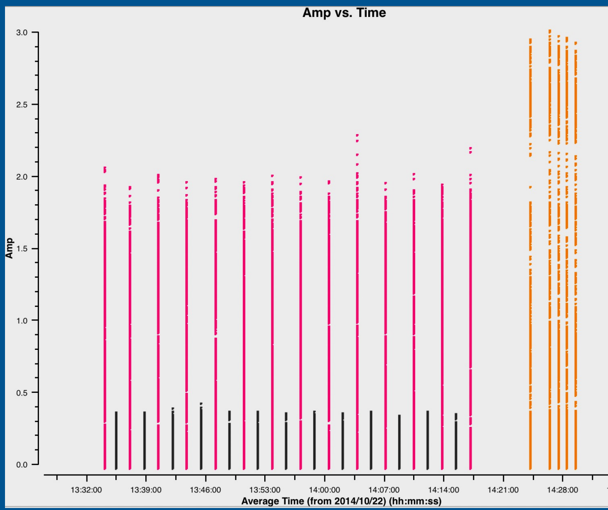
Can you tell if this is a phase referencing experiment or not?
How many sources are there?
What is the frequency setup?
How many antennas participated?

```
Fields: 5
  ID   Code Name                RA               Decl           Epoch        nRows
  0         J1640+3946          16:40:29.632770 +39.46.46.02836 J2000       254400
  1         3C345               16:42:58.809965 +39.48.36.99402 J2000       383520
  2         J1849+3024          18:49:20.103406 +30.24.14.23712 J2000       276480
  3         1848+283            18:50:27.589825 +28.25.13.15523 J2000       388800
  4         2023+336            20:25:10.842114 +33.43.00.21435 J2000       542880
Spectral Windows:  (8 unique spectral windows and 1 unique polarization setups)
  SpwID  Name   #Chans   Frame   Ch0(MHz)  ChanWid(kHz)  TotBW(kHz) CtrFreq(MHz)  Corrs
  0      none      32    GEO     4926.990       500.000     16000.0   4934.7400   RR  RL  LR  LL
  1      none      32    GEO     4942.490       500.000     16000.0   4950.2400   RR  RL  LR  LL
  2      none      32    GEO     4958.990       500.000     16000.0   4966.7400   RR  RL  LR  LL
  3      none      32    GEO     4974.490       500.000     16000.0   4982.2400   RR  RL  LR  LL
  4      none      32    GEO     4990.990       500.000     16000.0   4998.7400   RR  RL  LR  LL
  5      none      32    GEO     5006.490       500.000     16000.0   5014.2400   RR  RL  LR  LL
  6      none      32    GEO     5022.990       500.000     16000.0   5030.7400   RR  RL  LR  LL
  7      none      32    GEO     5038.490       500.000     16000.0   5046.2400   RR  RL  LR  LL
```

```
Antennas: 12:
  ID   Name  Station   Diam.    Long.         Lat.

  0    EF    EF        0.0  m   +006.53.01.0  +50.20.09.1
  1    WB    WB        0.0  m   +006.38.00.0  +52.43.48.0
  2    JB    JB        0.0  m   -002.18.30.9  +53.03.06.6
  3    ON    ON        0.0  m   +011.55.04.0  +57.13.05.3
  4    NT    NT        0.0  m   +014.59.20.6  +36.41.29.4
  5    TR    TR        0.0  m   +018.33.50.6  +52.54.37.8
  6    SV    SV        0.0  m   +029.46.55.0  +60.22.02.0
  7    ZC    ZC        0.0  m   +041.33.54.6  +43.35.44.2
  8    BD    BD        0.0  m   +102.14.02.1  +51.34.59.0
  9    SH    SH        0.0  m   +121.11.58.8  +30.55.45.2
  10   HH    HH        0.0  m   +027.41.07.4  -25.44.20.1
  11   YS    YS        0.0  m   -003.05.12.7  +40.20.05.0
~
~
```

```
plotms(vis='n14c3_TY_GC_flags.ms/', <…>, …)
```

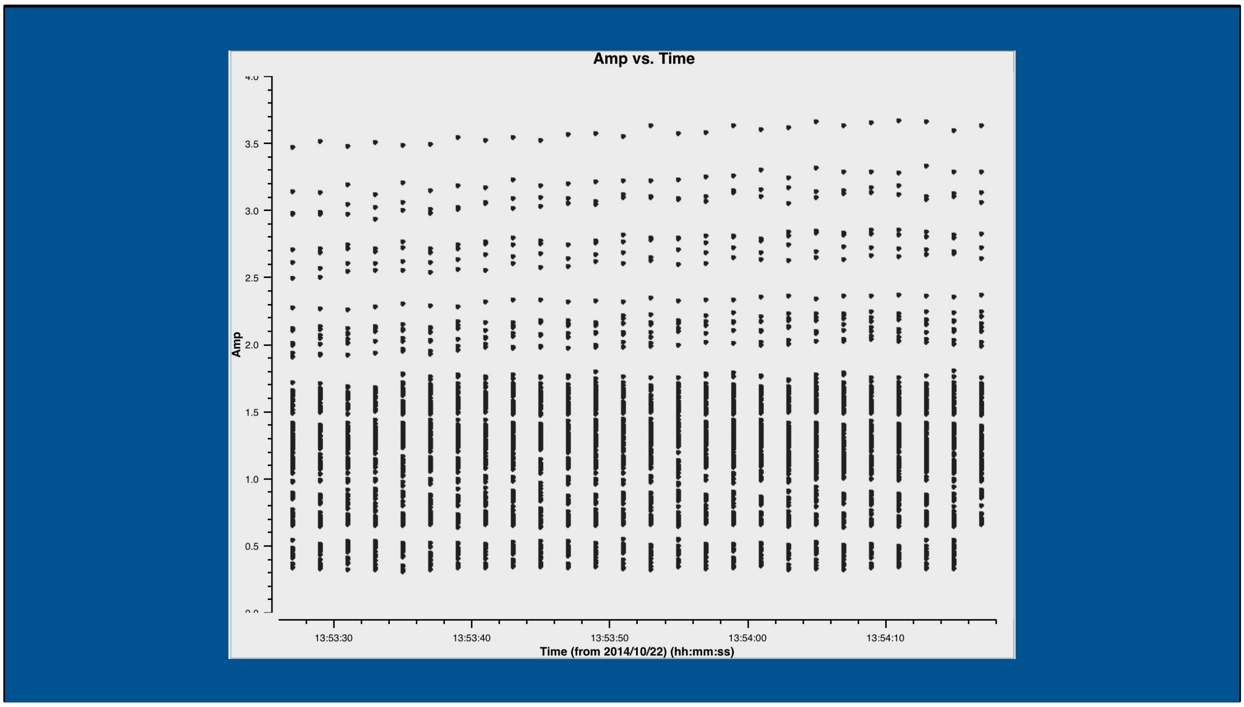


The plot clearly shows this is a phase referencing experiment, as shown in the lecture. In this case the phase calibrator is also bright enough to correct for the instrumental delay.
This experiment has two pairs of target-phasecal, in principle one instrumental delay correction is enough for the entire experiment, but you can also split the experiment into two and handle each separately.

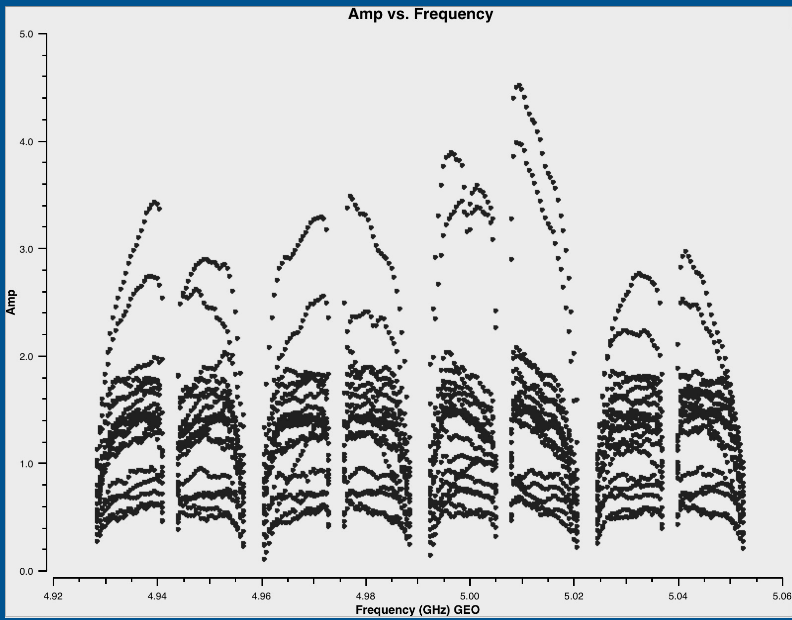
2. Selecting a calibration scan

- Bright source
- Middle of the observation
- Single scan for the instrumental delay correction
- Verify:
  - All antennas present
  - Stable phases in time/freq
  - No weird stuff in amplitudes

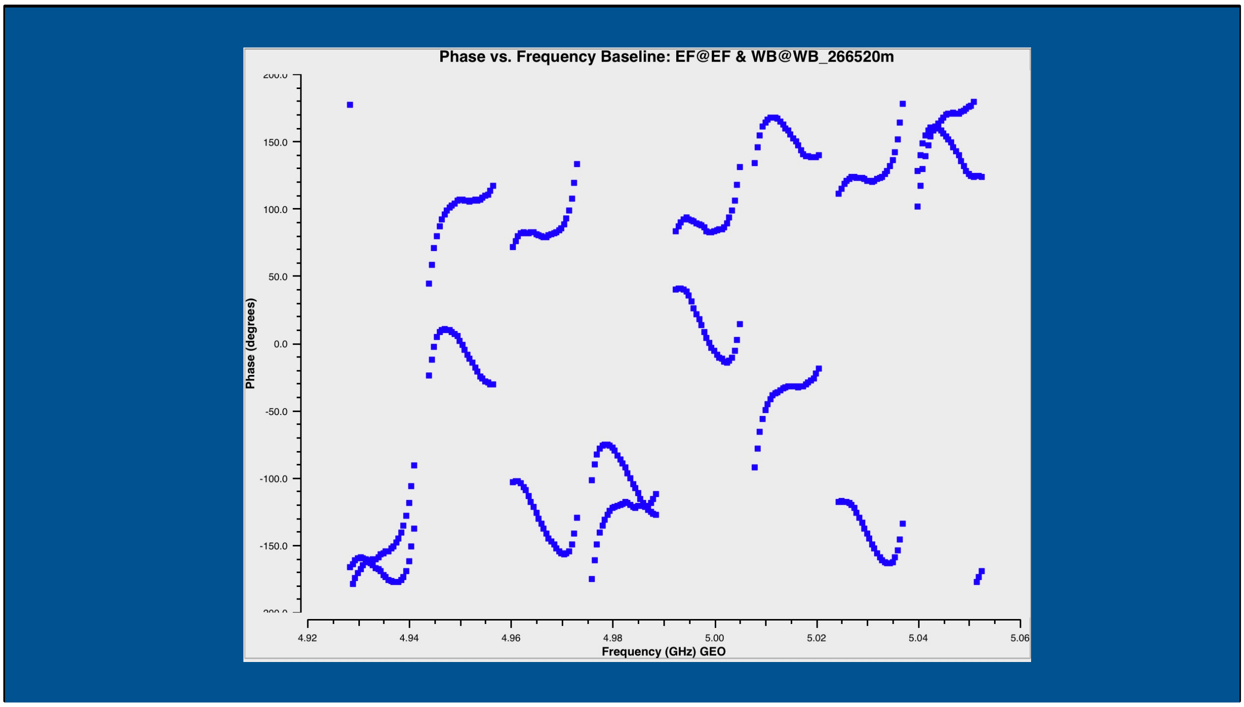
For the instrumental delay calibration we need a bright point source.
What to do if your phasecal is resolved? In CASA you can calibrate using a source model.



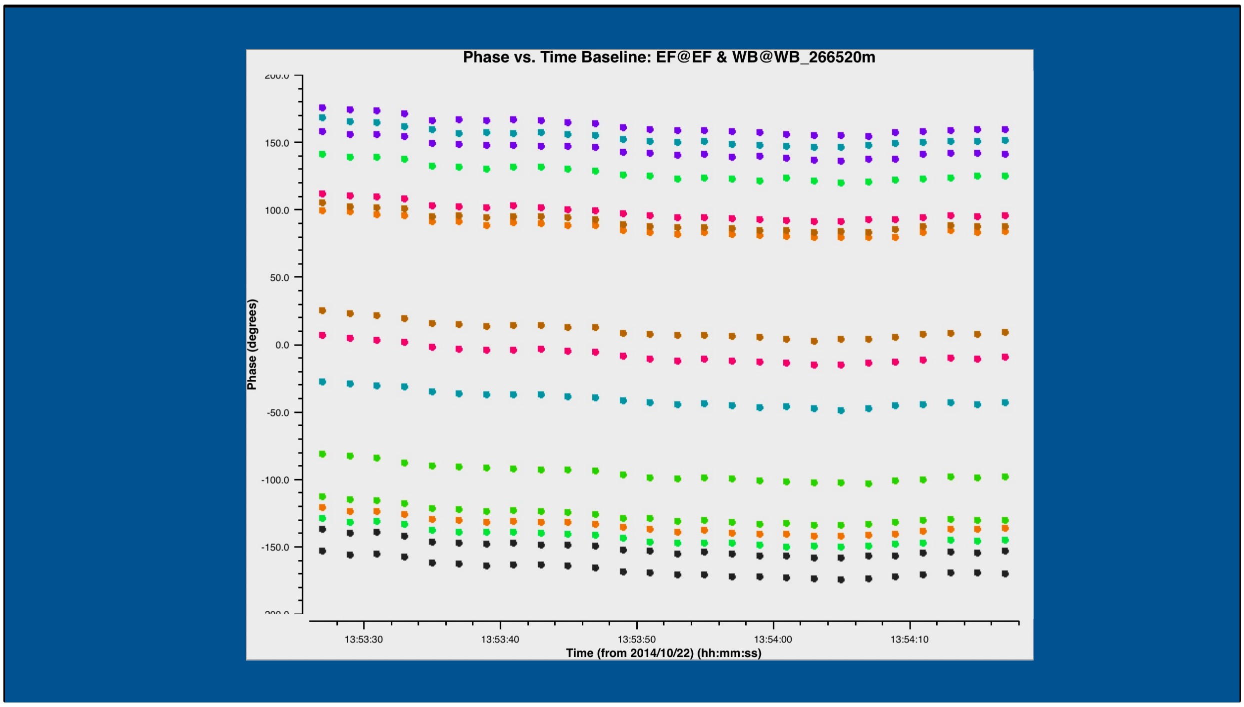
Data for EF baselines only: this looks OK, though there is a selection of points that are clearly higher than the average in each integration. This is due to the JodrellBank (JB) station.



Again the higher data are the JB observations, but they look OK. So scaling the amplitudes in the self-calibration may fix this.



The phases on a single baseline look good. The spectral windows are clearly offset from each other, but within each the phase is continuous.



Colour is the different spectral windows, two per polarization. Phases are quite stable within this scan, so it is good for fringe fitting.

3. Fringe fitting

CASA task: `fringefit`

- Requires:
  - Single scan on bright target
  - Select middle timerange of the scan, ~1-2 minutes
  - Select a reference antenna
  - Always set parang = True for VLBI
  - Instrumental delay: zerorates=true

Fringe fitting is done to calculate the instrumental delay also known as single band delay

```
# fringefit -- Fringe fit delay and rates
vis                = ''                        # Name of input visibility f
caltable           = ''                        # Name of output gain calibr
field              = ''                        # Select field using field i
spw                = ''                        # Select spectral window/cha
intent             = ''                        # Select observing intent
selectdata         = True                      # Other data selection param
   timerange       = ''                        # Select data based on time
   antenna         = ''                        # Select data based on anten
   scan            = ''                        # Scan number range
   observation     = ''                        # Select by observation ID(s
   msselect        = ''                        # Optional complex data sele
solint             = 'inf'                     # Solution interval: egs. 'i
combine            = ''                        # Data axes which to combine
refant             = ''                        # Reference antenna name(s)
minsnr             = 3.0                       # Reject solutions below thi
zerorates          = False                     # Zero delay-rates in soluti
globalsolve        = True                      # Refine estimates of delay
niter              = 100                       # Maximum number of iteratio
delaywindow        = []                        # Constrain FFT delay search
ratewindow         = []                        # Constrain FFT rate search
append             = False                     # Append solutions to the (e
corrdepflags       = False                     # Respect correlation-depend
docallib           = False                     # Use callib or traditional
   gaintable       = []                        # Gain calibration table(s)
   gainfield       = []                        # Select a subset of calibra
   interp          = []                        # Temporal interpolation for
   spwmap          = []                        # Spectral window mappings t
paramactive        = []                        # Control which parameters a
parang             = False                     # Apply parallactic angle co
```

The solutions calculated by the task will be written to a calibration table
Important parameters here are the gaintable, gainfield, zerorates and parang.
To get the highest quality data, take a short interval in the middle of a scan (1 minute or so)
As interferometry does not have an absolute zero point for phase, we need a reference antenna too

Selecting a reference antenna

- For EVN: use Effelsberg (EF)
  - Biggest dish
  - Well behaved meta-data
  - Central in the array
- The parameter accepts a string, e.g.
  `refant = "EF,TR,ON"`
- If not sure: use CASA task `plotants`

Selecting a reference antenna for the EVN: we usually pick Effelsberg (EF). We are not limited to one reference antenna, the parameter accepts a string and goes through the list in the order provided. So in this example, EF is chosen first, if there is no data for EF the task will try Torun (TR). And if there is no data for TR either it will try Onsala (ON).

On-the-fly calibration

- `docallib = False`
- `gaintable=['n14c3_ERIS.tsys','n14c3_ERIS.gcal']`
- `interp=['nearest,nearest','nearest']`

An important aspect of CASA is that each task does on-the-fly calibration. Before fringe fitting the system temperature and gain curve corrections are applied, then the fringefit task will work with those calibrated data to calculate the phase and delay solutions. The docallib parameter refers to the use of calibration libraries, which we do not use, so it can be set to False. The order of the gaintables is irrelevant, CASA will always apply them in the correct order.

All fringefit parameters

```
fringefit(vis='n14c3_TY_GC_flags.ms/',
    caltable='n14c3_ERIS.sbd',
    timerange='13:53:20.0~13:54:20.0',
    solint='inf',
    refant='EF',
    minsnr=50,
    zerorates=True,
    corrdepflags=True,
    gaintable=['n14c3_ERIS.gcal', n14c3_ERIS.tsys'],
    interp=['nearest','nearest,nearest'],
    parang=True
)
```

These are all the settings you want for the instrumental delay correction.
Caltable: can be any name of your choice. I use the experiment name, and an extension that tells me what type of calibration it is. You can think up your own system, whatever works to ensure you will know what's in the table in a year from now.
Timerange: this is the middle minute of scan, the data we just looked at
Solint: as we want all the data for a single solution, we set the interval to 'inf'
Minsnr: for the instrumental delay you want really good solutions, and you should have a bright source. A value of 50 ensures that the data are sufficient quality, but if you find antennas having values between 50 and 100 you may want to check what's happening there.
Zerorates: we only need a delay solutions, so we force the rates to be set to 0.
Corrdepflags: in case a station has only a single polarization, this will keep that data if set to True. When set to False, single-polarization data will be flagged entirely and the data will be lost for further calibration.
Parang: always, always, always use parallactic angle correction for VLBI calibration

4. Inspect the results

```
2022-09-19 13:23:28  WARN          …ingefit::::  MS obs=0,fld=0,spw=0,ant=12 cannot be calibrated
2022-09-19 13:23:28  WARN          …ingefit::::  MS obs=0,fld=0,spw=1,ant=12 cannot be calibrated
2022-09-19 13:23:28  WARN          …ingefit::::  MS obs=0,fld=0,spw=2,ant=12 cannot be calibrated
2022-09-19 13:23:28  WARN          …ingefit::::  MS obs=0,fld=0,spw=3,ant=12 cannot be calibrated
2022-09-19 13:23:28  WARN          …ingefit::::  MS obs=0,fld=0,spw=4,ant=12 cannot be calibrated
2022-09-19 13:23:29  WARN          …ingefit::::  MS obs=0,fld=0,spw=5,ant=12 cannot be calibrated
2022-09-19 13:23:29  WARN          …ingefit::::  MS obs=0,fld=0,spw=6,ant=12 cannot be calibrated
2022-09-19 13:23:29  WARN          …ingefit::::  MS obs=0,fld=0,spw=7,ant=12 cannot be calibrated
```

This is normal, don’t worry!

When there is no calibration information for a visibility, you will get a warning. This will happen frequently and is normal.

4. Inspect the results

```
Using reference antenna 0
sPok [8, 13]
Antenna 1 correlation 0 has (FFT) SNR of 1609.38.
Antenna 2 correlation 0 has (FFT) SNR of 1926.22.
Antenna 3 correlation 0 has (FFT) SNR of 5533.
Antenna 4 correlation 0 has (FFT) SNR of 3867.23.
Antenna 5 correlation 0 has (FFT) SNR of 5968.37.
Antenna 6 correlation 0 has (FFT) SNR of 0 below threshold (50).
Antenna 7 correlation 0 has (FFT) SNR of 7999.78.
Antenna 8 correlation 0 has (FFT) SNR of 6248.67.
Antenna 9 correlation 0 has (FFT) SNR of 885.085.
Antenna 10 correlation 0 has (FFT) SNR of 0 below threshold (50).
Antenna 11 correlation 0 has (FFT) SNR of 10894.2.
Antenna 1 correlation 1 has (FFT) SNR of 1596.69.
Antenna 2 correlation 1 has (FFT) SNR of 2517.05.
Antenna 3 correlation 1 has (FFT) SNR of 4479.48.
Antenna 4 correlation 1 has (FFT) SNR of 4135.53.
Antenna 5 correlation 1 has (FFT) SNR of 17669.6.
Antenna 6 correlation 1 has (FFT) SNR of 9734.02.
Antenna 7 correlation 1 has (FFT) SNR of 6020.06.
Antenna 8 correlation 1 has (FFT) SNR of 9241.53.
Antenna 9 correlation 1 has (FFT) SNR of 901.532.
Antenna 10 correlation 1 has (FFT) SNR of 1171.34.
Antenna 11 correlation 1 has (FFT) SNR of 8385.43.
Starting least squares optimization.
Least squares complete for correlation 0 after 3 iterations.
Least squares complete for correlation 1 after 3 iterations.
Zeroing delay rates in calibration table.
```

In the logger, check the SNR for the solutions. These will be provided for each polarization, each spectral window and each antenna. These values look fine, in the case of antenna 6 and 10 for correlation 0 there are no data, so SNR=0. If an antenna has very low SNR, you want to inspect the data further.
Another value to check is the number of iterations, for a bright source and instrumental delay calibration this number should be small, a few to 10-ish iterations is OK.

4. Inspect the results

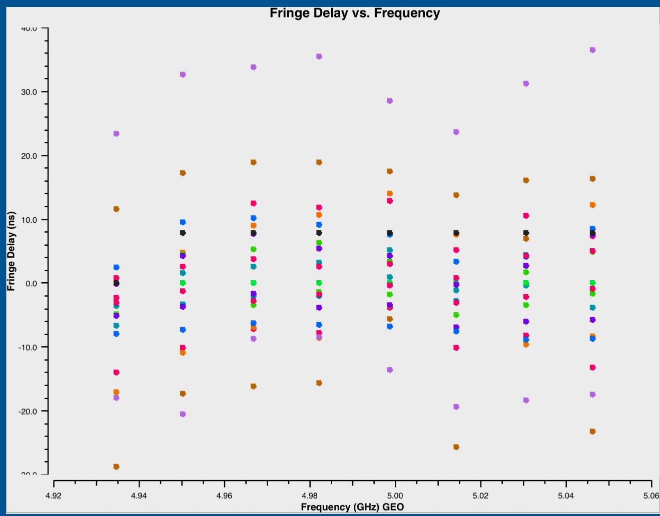
```
Finished solving.
Calibration solve statistics per spw:  (expected/attempted/succeeded):
  Spw 0: 1/1/1
  Spw 1: 1/1/1
  Spw 2: 1/1/1
  Spw 3: 1/1/1
  Spw 4: 1/1/1
  Spw 5: 1/1/1
  Spw 6: 1/1/1
  Spw 7: 1/1/1
Task fringefit complete. Start time: 2022-09-19 15:23:25.938173 End tim
##### End Task: fringefit            #####
```

When fringefit is done it produces a success rate. It expected one solution per spectral window, attemted one and succeeded. This is the case for all spectral windows, which is good.

Plot the calibration table

CASA task `plotms`

Instead of MS, use calibration table as `vis`



First inspect the CASA logger output, then plot the solutions.
The delay is typically in the order of nano-seconds. The colours are the stations, and you can see that some stations have somewhat higher values, but these are consistently high for all spectral windows.

Plot the calibration table

```
plotms(vis='n14c3_ERIS.sbd',
     xaxis='freq',
     yaxis='delay',
     coloraxis='antenna1')
```

Plotms can not only handle visibility data, it also plots calibration tables.

5. Apply the calibration

CASA task `applycal`

Creates CORRECTED_DATA column in your MS

- Add all available calibration tables
- Set the interpolation
- Set parang=True

- To clean-up a calibration: task `clearcal`

As a final check, apply the solutions to the data and verify that the calibrated phases for the selected timerange are corrected: they should be flat and around zero degrees for all baselines.

5. Apply the calibration

```
applycal(vis='n14c3_TY_GC_flags.ms/',
gaintable=['n14c3_ERIS.gcal','n14c3_ERIS.tsys',
'n14c3_ERIS.sbd'],
interp=['nearest', 'nearest,nearest', 'nearest'],
parang=True,
flagbackup=False
)
```

Plot the calibrated phases

```
plotms(vis='n14c3_TY_GC_flags.ms',
     xaxis='freq',
     yaxis='phase',
     ydatacolumn='corrected',
     timerange='13:53:20~13:54:20',
     antenna='EF',
     coloraxis='baseline',
     correlation='LL',
     averagedata=True,
     avgtime='120'
)
```

This should give a plot with all the phases around 0 for all spectral windows. If you change 'corrected' to 'data' you can see the difference between calibrated and uncalibrated phases.