LOFAR CS1 IMAGING PIPELINE

LOFAR CORE STATION PROTOTYPE DATA PROCESSING PIPELINE

G.A. Renting, R.J. Nijboer

renting@astron.nl

1 Overview

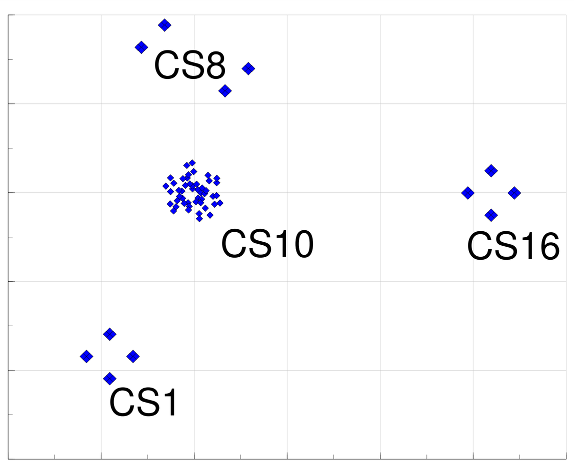
LOFAR

LOFAR is the Low Frequency Array radio telescope currently under construction in the Netherlands. It will consist of 77 stations of 96 High Band Antennas (HBA) and 96 Low Band Antennas (LBA) spread out over an area with a diameter of over 100 km. International cooperation will extend the area to cover over 350km baselines. The HBAs observe in the range from 120 to 250 Mhz, and the LBAs observe in the range from 30 to 90 Mhz. There will be a core of 32 stations and 5 arms with 9 stations each, which should become operational during 2008.

LOFAR station positions

Core Station 1

Core Station 1 is a nickname for a prototype consisting of the first four LOFAR stations, numbered CS1, CS8, CS10 and CS16. These are being installed in the field in september/ october 2006, one year in advance of the full LOFAR roll-out. The goal is to test prototypes of all components, connections and software of the final system.
CS10 contains 48 antennas and the other three stations contain 16 antennas each. For testing purposes these can be either operated as single stations, or subdived into microstations to increase the number of baselines.

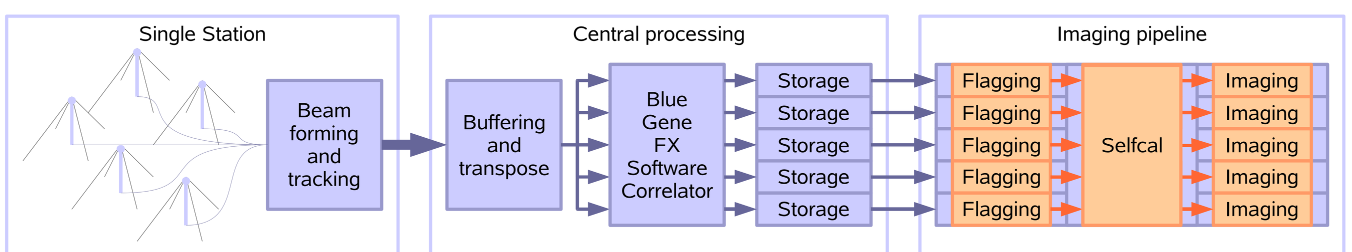


CS1 antenna positions

Data processing pipeline

The data from the antennas is processed in several steps shown in the graph below:

- Electronic beam forning and tracking at the station level to reduce the data from the individual antennas to a single datastream per station.
- Input buffering and transpose from station ordering to subband ordering at the input cluster.
- Correlation per subband using a dedicated 12.288 cores, 34.4 TFlop, IBM BlueGene/L.
- Storage of 1 second integration raw uv-data.
- Imaging pipeline.



Data processing pipeline

CS1 Imaging pipeline

For CS1 the imaging pipeline will not be iterative or contain any feedback mechanisms. It will consist of the following three basic blocks that are described in more detail on the rest of this presentation:

- Flagging module operating independently on each subband. (2)
- Self-calibration module operating in parallel on all subbands distributed across multiple processing nodes. (3)
- Imaging module operating on all subbands separately. (4)

2 Flagging module

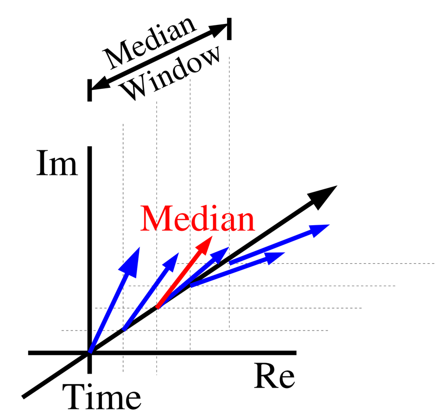
An important problem at these low frequencies and at this geographical location is human generated Radio Frequency Interference (RFI). A module has been developed to flag uv-data points for which RFI dominates to a level where the inspected data can not be used in further processing.
The algorithm works on the assumption that RFI will rotate quickly in the complex plane compared to celestial signals. A complex median value is computed over a window in time per baseline, spectral window, channel and polarization. If the distance in the complex plane to this median is above a certain treshold the data point is flagged for exclusion.

$$\mathit{flag\ inspected\ point} = \left|\mathit{inspected\ point} - \mathit{complex\ median}\right| < \mathit{baseline\ length\ dependent\ threshold}$$

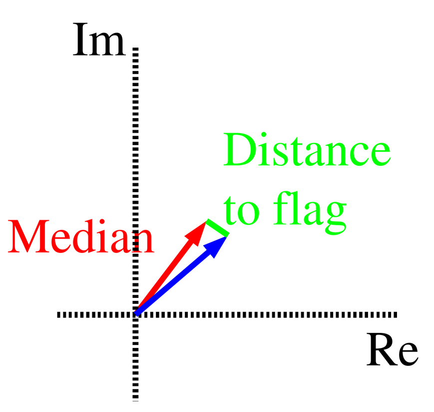
The threshold is a baseline length dependent multiple of a theoretical noise estimate based on integration time and channel bandwidth.

$$Noise_{Estimate} = \frac{k}{\sqrt{t_{integration} \cdot Bandwidth_{channel}}}$$

The baseline length dependency is used since observing a strong source will produce fringes in the data that rotate much faster on longer baselines.



Complex median

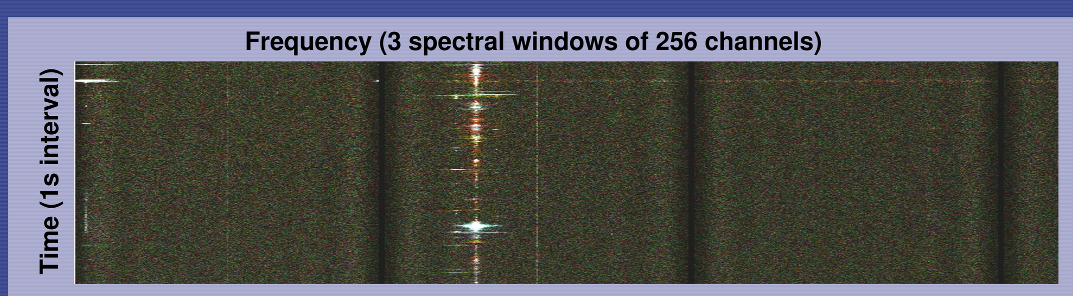


Flagging criterion

Example of flagging on CS1 data

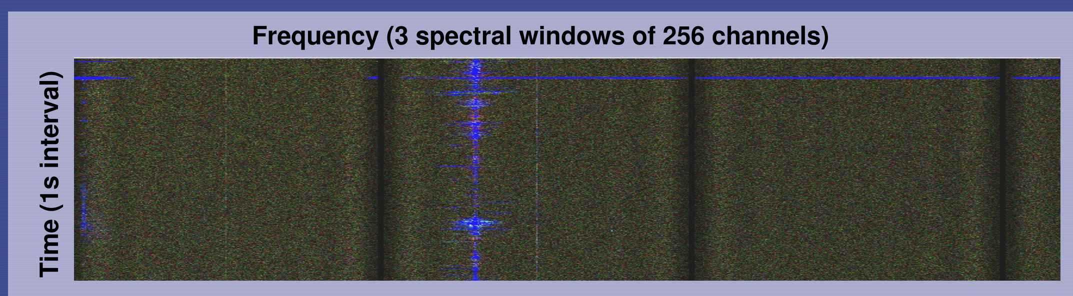
The two figures on the right show visibility data of a 60-61 Mhz CS1 measurement.

The top image shows a lot of interference in the second spectral window.



Unflagged data, intensity marks amplitude, color marks phase

The bottom image shows almost all interference is marked as flagged.



Flagged data, blue pixels mark flagged data points

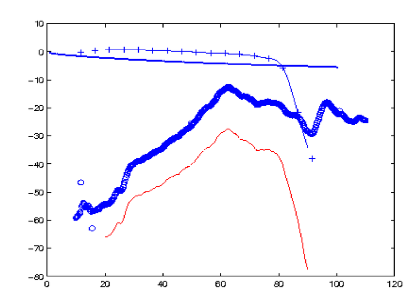
3 Self-calibration module

The CS1 selfcalibration module is called Black Board Selfcal (BBS) and is based on the Measurement Equation formalism. The Measurement Equation gives the visibilities as shown in the following formula:

$$V_{ij} = \sum_k \int df \int dt \,(J_{ik} \otimes J^*_{jk})\, S\, \vec{I}_k$$

BBS can solve for parameters that make up different Jones matrix (J) contributions, like Bandpass (B), Electronic Gain (G), Beamshape (E), Receptor orientation (P), Ionopheric Phase (I), Farady Rotation (F), and Fourier kernel (K) as given in the following equation:

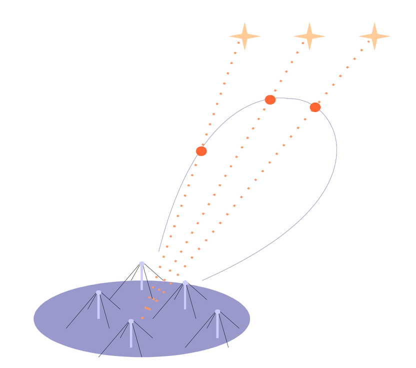
$$J_{ik} = B_i G_i E_{ik} P_{ik} I_{ik} F_{ik} K_{ik}$$



Constructing a bandpass profile

For CS1 only a very simple selfcal strategy is implemented to test the current BBS selfcal software:

- For each timeslot a combination of measured bandpass contributions is fitted to the data. The profiles used are continuous over the entire observing window, and not calculated separately for each subband. In effect this is taking the B as a fixed function of frequency and solving for G in time.
- Solving for the Jones matrices E, P, I, F, K as a single matrix J for a few different directions in the sky.
- Fitting a beamshape model to the calibrated directions.
- Substracting the detected sources from the data.
- Correct the residual data for a single point in the sky.

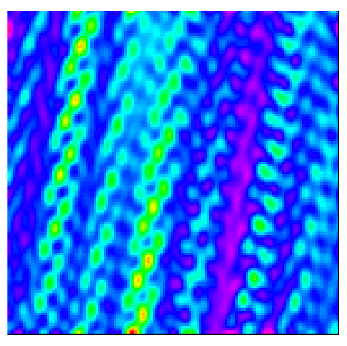


Fitting beamshape

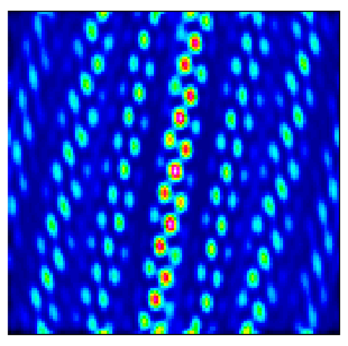
4 Imaging module

The CS1 imaging module is based on the AIPS++ imager, and is used to process each subband to create frequency image cubes and continuum images.

One of the first images is shown below. As CS1 is still under construction, only 8 antennas are currently fully connected resulting in a poor point spread function and thus image quality.

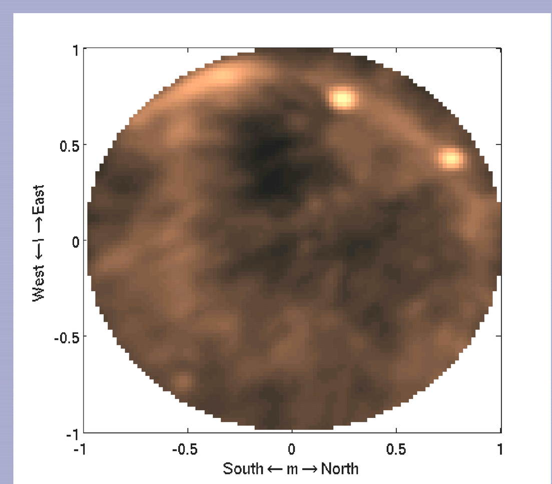


CS1 Imager result



Point Spread function

The All sky image below shows an image created by data obtained directly from a single station. It contains 43 subbands from 48 antennas around 57 Mhz and is imaged with a custom DFT imager.



All sky image showing CasA and CygA.

Image courtesy of Stefan Wijnholds.