VIRAC automated Single baseline interferometre data processing

Jānis Šteinbergs[1], Karina Šķirmante[1], Ross Burns[1, 2]

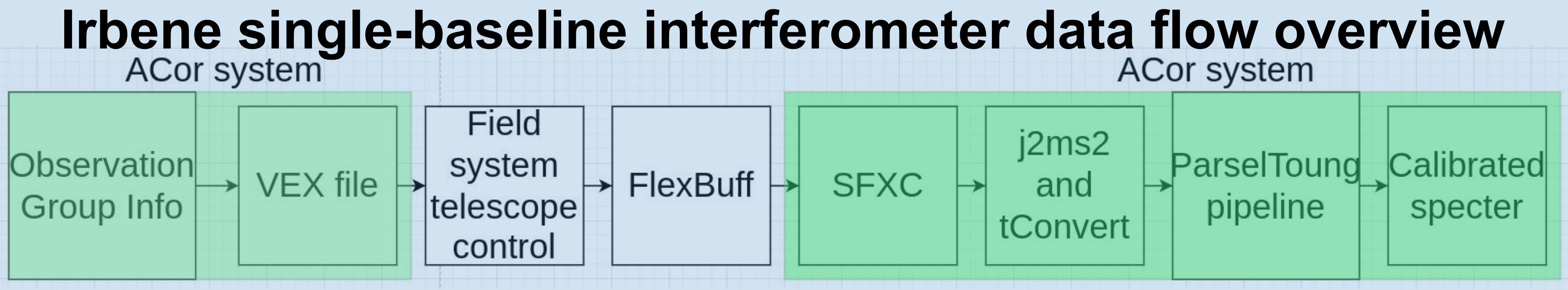
1. Engineering Research Institute Ventspils International Radio Astronomy Centre of Ventspils University of Applied Sciences
2. RIKEN Cluster for Pioneering Research, 2-1 Hirosawa, Wako-shi, Saitama, 351-0198, Japan

16th European VLBI Network Symposium and Users' Meeting, 2-6 September 2024

Abstract

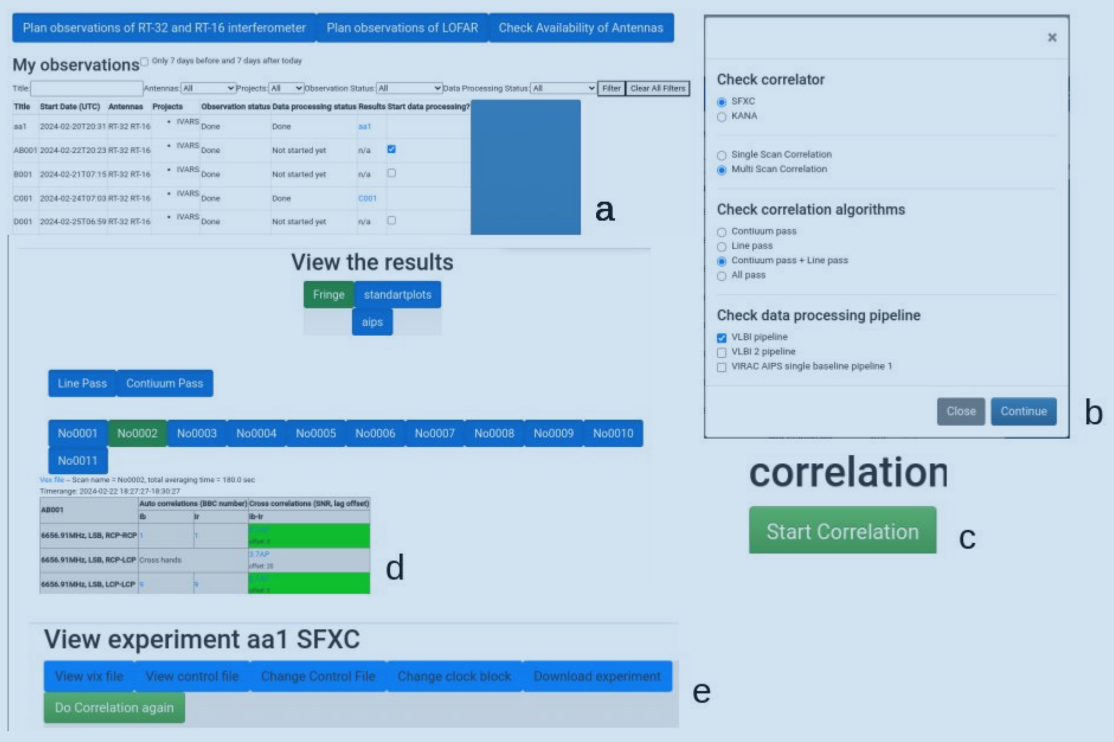
High-mass star formation has entered the discipline of time-domain astronomy. These bursts of accretion cause variations in the radio continuum emission and radio frequency maser emission emitted by protostars. The associated rising demand for high-cadence observations calls for the development of observational facilities that can effectively monitor the radio frequency continuum and maser emission in a manner that provides high detection sensitivity and can be highly automated. With highly automated observation and data processing infrastructure, in order to monitor a sample of 30 high-mass protostars. In this poster we describe the automation tools developed as part of the IVARS project.

Irbene single-baseline interferometer data flow overview



The VIRAC data centre consists of 1) Flexbuffs - for initial data storage after observation, 2) LOFAR data server - data storage for data processing and 3) high-performance computer for data processing. Within the VIRAC data center, the ACor system functions are: 1) generate key file and creates VEX file for specific observation and group, 2) send VEX file to telescope operators, 3) send data from flexbuff to LOFAR data server, 4) generate files for SFXC correlation, 5) run correlations and create measurement-set and FITS files, 6) run the ParselTounge pipeline.

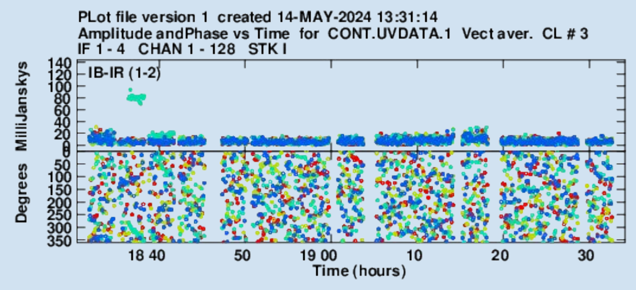
The ACor view of observation data processing



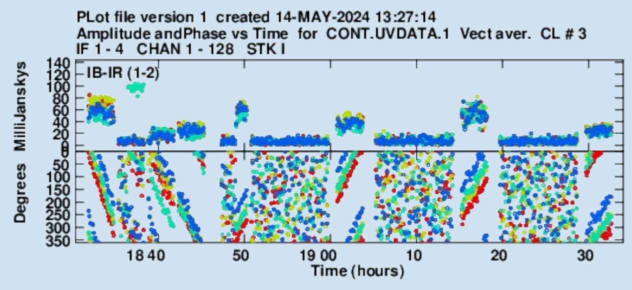
a) View of observations and status,
b) Data processing parameters choose view
c) View start correlation
d) Data processing result view
e) Experiments view

Automation of Correlation

- The ACor system allows processing of two kinds of observation.
  - Data processing with single a scan correlation - to experiment with correlation parameters.
  - Data processing with multi-scan correlation - to run a full correlation pass of line or continuum, or both.
- Multi scan correlation run includes all observed scans. Multi-scan correlation uses the following correlation parameters:
  - Two second integration time,
  - In the continuum pass, all channels are correlated with 128 FFT points,
  - The line-only channel containing the maser signal is correlated with 4096 FFT points.
  - These parameters are well suited for spectral resolution in maser observations using a 8 MHz bandwidth
- For multi-scan correlation an automatic clock search is executed. his is done by parsing the key file to identify the fringe finder source(s)
  - After that it is checked if the scan witch contains the fringe finder has raw data on file.
  - If none of the fringe finder scans have raw data files a clock search is not done.
  - The clock search is done with 1024 FFT points, 2 s integration time and for all channels.
  - This process is done in 5 iterations after the first iteration fringe mean offset value is found.
  - If the standard deviation of the fringe offset for all channels is larger than 2, the clock search is stopped.
  - In the next four iterations the identified fringe mean offset is subtracted and added to RT-32 GPS offset and RT-16 GPS offset.
  - After all iteration from all four GPS offset changes, the version where mean of fringe offset is lowest is chosen as the appropriate value for the full correlation pass.



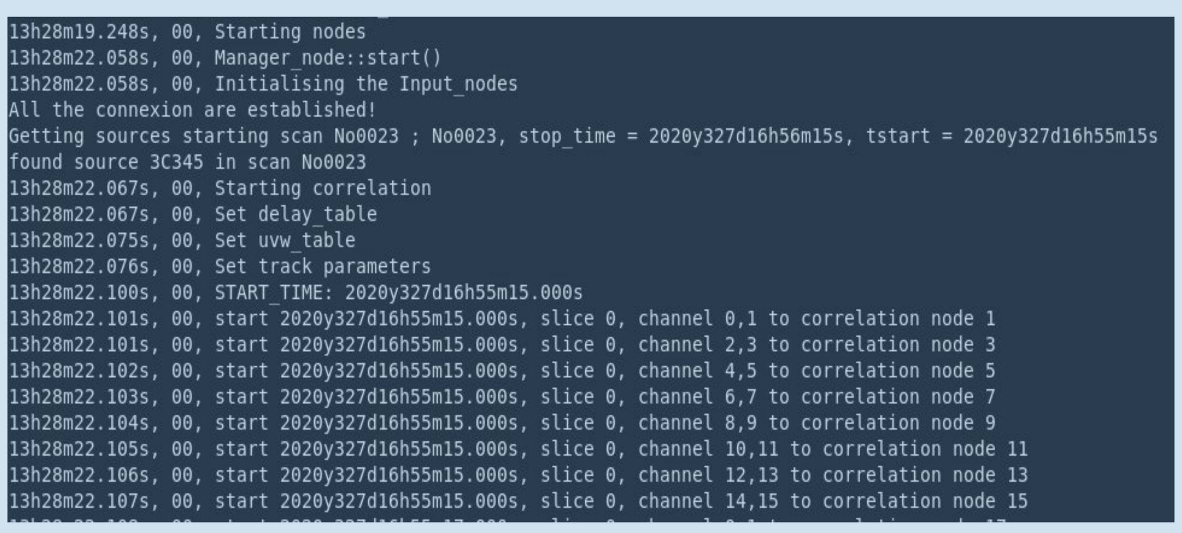
Data before clock search



Data after clock search

VIRAC HPC

- **The second generation** - each of 9 nodes has 2 X Intel Gold 24 cores processors; RAM 384 GB per node;10Gbps and 40Gbs Infiniband networks; 240GB SSD
- **The first generation** - each of 30 nodes has 2 processors with 8 cores for each, RAM 128 GB per each, 600 GB HDD per item;10Gbps and 40Gbs Infiniband networks.



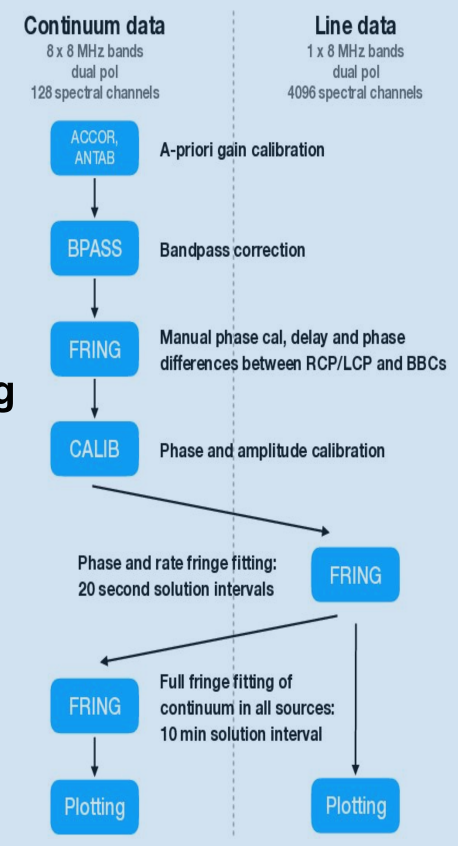
```
13h28m19.248s, 00, Starting nodes
13h28m22.058s, 00, Manager_node::start()
13h28m22.058s, 00, Initialising the Input_nodes
All the connexion are established!
Getting sources starting scan No0023 ; No0023, stop_time = 2020y327d16h56m15s, tstart = 2020y327d16h55m15s
found source 3C345 in scan No0023
13h28m22.067s, 00, Starting correlation
13h28m22.067s, 00, Set delay_table
13h28m22.075s, 00, Set uvw_table
13h28m22.076s, 00, Set track parameters
13h28m22.100s, 00, START_TIME: 2020y327d16h55m15.000s
13h28m22.101s, 00, start 2020y327d16h55m15.000s, slice 0, channel 0,1 to correlation node 1
13h28m22.101s, 00, start 2020y327d16h55m15.000s, slice 0, channel 2,3 to correlation node 3
13h28m22.102s, 00, start 2020y327d16h55m15.000s, slice 0, channel 4,5 to correlation node 5
13h28m22.103s, 00, start 2020y327d16h55m15.000s, slice 0, channel 6,7 to correlation node 7
13h28m22.104s, 00, start 2020y327d16h55m15.000s, slice 0, channel 8,9 to correlation node 9
13h28m22.105s, 00, start 2020y327d16h55m15.000s, slice 0, channel 10,11 to correlation node 11
13h28m22.106s, 00, start 2020y327d16h55m15.000s, slice 0, channel 12,13 to correlation node 13
13h28m22.107s, 00, start 2020y327d16h55m15.000s, slice 0, channel 14,15 to correlation node 15
```

SFXC run

VIRAC HPC

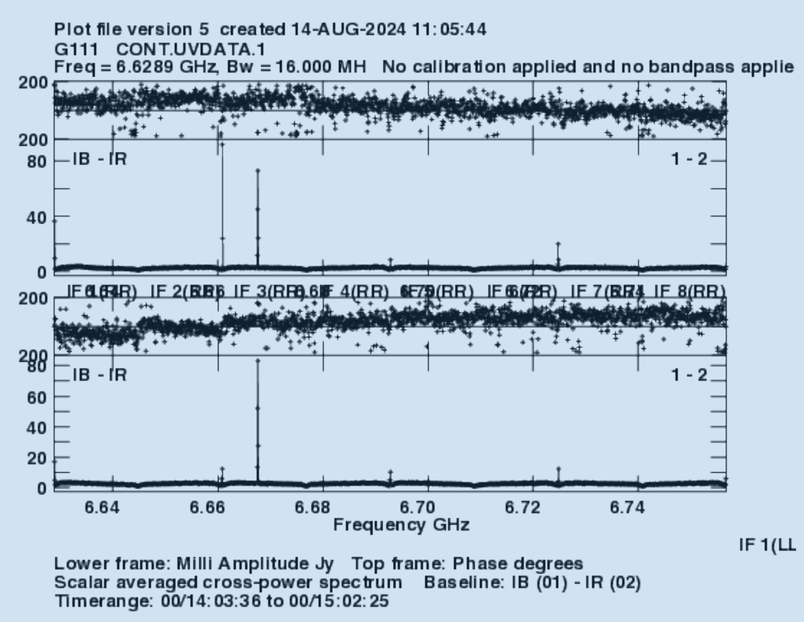
Automation of data reduction

- Processing of astronomical data was standardised with the use of an automated data reduction pipeline written in ParselTongue which is a python interface to the Astronomical Image Processing Software (AIPS).
- The line and continuum data sets are first loaded
- Beginning with the continuum data, corrections for the losses during digital sampling are applied
- Then a-priori gain calibration tables derived from noise diode temperature measurements conducted during observations are used to calibrate the flux density of visibilities.
- Bandpass corrections for all targets are made based on the observed bandpass shapes of continuum calibrators
- Three stages of fringe fitting and integrations are performed, including instrumental 'manual phase-cal', group delay, and source continuum fringe fitting, each with specific purposes and parameters. Solutions are applied to all sources.
- The peak channel of maser emission is then used to calibrate rapid phase fluctuations and the solutions of this fringe fitting are applied back on to the continuum emission of maser sources in order to enable long-duration coherent integration, aiming to detect the continuum emission associated with maser targets.
- Finally the spectra and visibility plots of all sources are output and the integrated radio continuum flux densities of all sources determined.

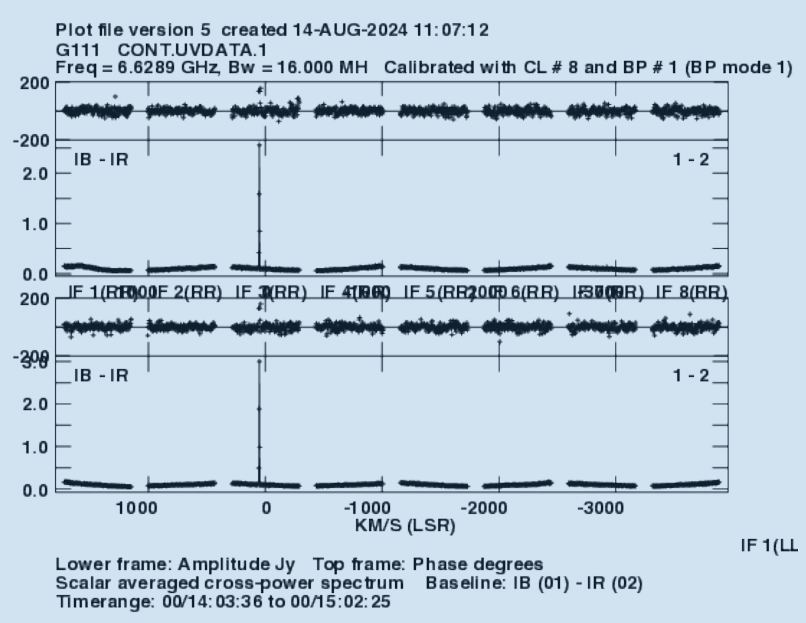


Visual of the main stages of the automated AIPS data processing carried out by the ParselTongue pipeline

GitHub Link to pipeline

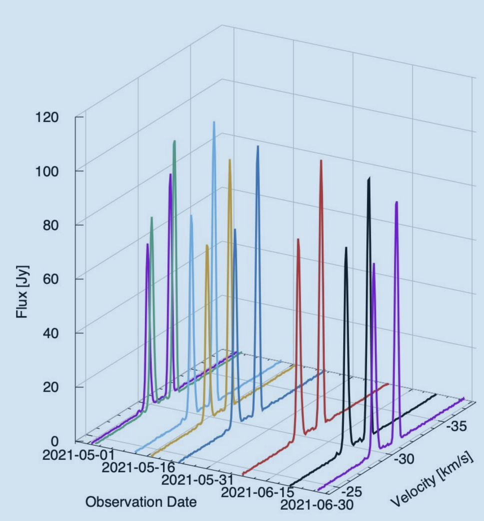


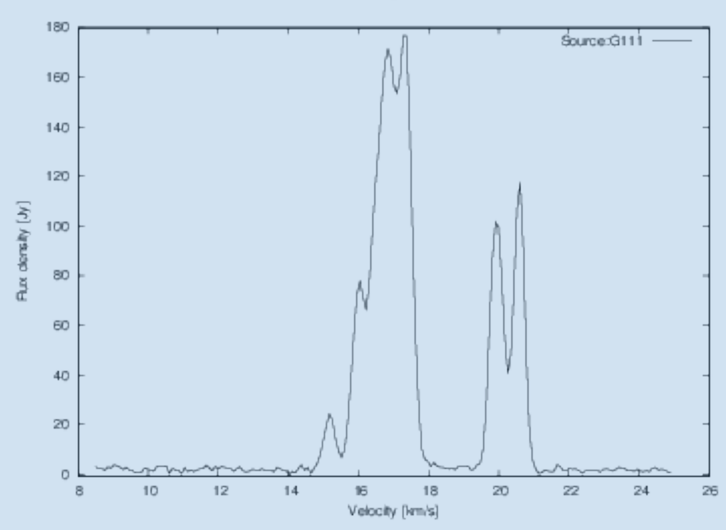
Before calibration



After calibration

First monitoring results





GitHub Link to ISBI monitoritng

Methanol maser emission from the same observation and target, from the high-resolution correlator pass

Methanol maser and continuum emission in high-mass protostar G085.410+0.003 from pilot observations from this project

Conclusions

- We have created automatic data processing and calibration syteme. Results can be seen in this poster.
- This allow us to conduct observations of he 800 m single-baseline interferometer at Irbene Radio Observator with very low usage of human resources. And it make complicated processes executable for every one.
- In future we will work on VLBI simulations, that will allow us better understand errors in data and improve calibration.

Acknowledgements

This publication has received funding from the Latvian Council of Science project "A single-baseline radio interferometer in a new age of transient astrophysics (IVARS)" (lzp-2022/1-0083).