Theme 3: Imaging and Non-imaging Algorithms and Software

During the workshop this group explored three main areas:

- The processing of SKA scale datasets with multiple software packages.
- Introduction to wsclean and ASKAPSoft.
- The application of novel time-domain analysis schemes to SKA precursor observations.

SKA-Scale Simulations and Imaging

One of existing projects between ICRAR and SHAO is to investigate SKA-Scale data processing. The project will simulate the complete pipeline from correlator through data ingest and imaging. During the workshop we generated the simulated datasets and utilised a number of packages to produce imaging data products.

The Simulations

The simulation uses the Murchison Widefield Array GLEAM image as a base, and uses 128,000 individual components, each with a physical extent, an intensity and a spectral index, to represent the image. We had limited processing power available and therefore could only simulate a few spectral channels. This work will continue on TianHe and will produce the full spectral range.

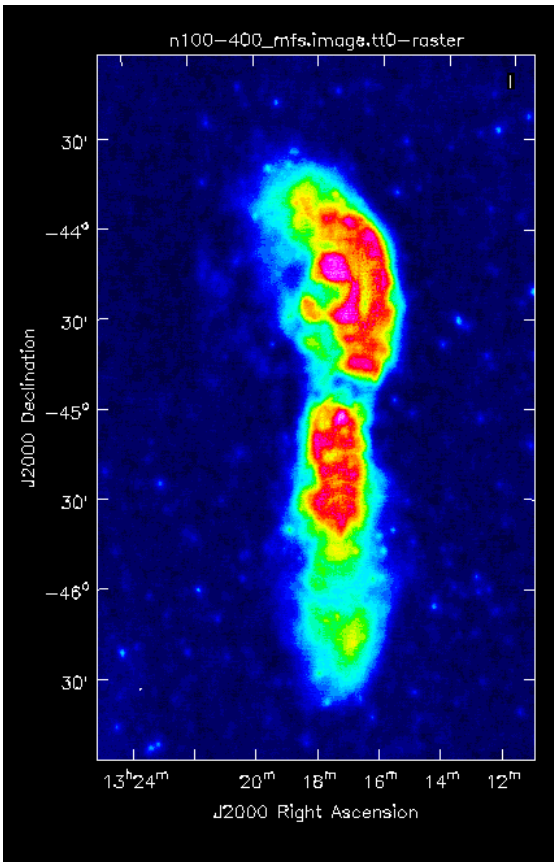

Figure 1 Total Intensity image of the simulation of Centarus A produced during the workshop

The Imaging

During the workshop we made significant progress imaging the simulations that we made during the meeting (see Figure 1).
We used using a number of different packages, the different packages all use different imaging and deconvolution schemes, and are in common use throughout the SKA precursor and wider astronomical community – all the packages were brought to the workshop and distributed using Docker images and all processing presented here was performed locally.

We compared the following processing tools:

- wsclean,
- casapy,
- ASKAPSoft.

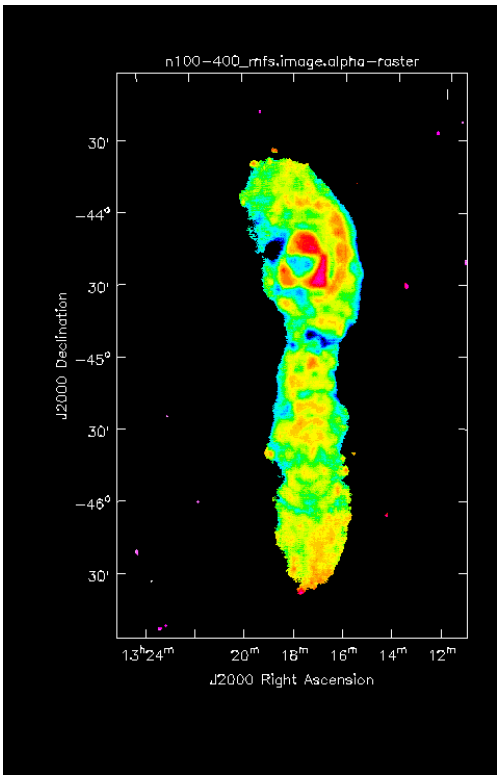

Figure 2 An image of the Centaurus A simulation in spectral index.

And generated the following products:

- Channel averaged maps
- Taylor term maps
- Spectral index maps

The spectral index map obtained was unfortunately very coarse as we were only able to perform this analysis on a small number of channels. But Figure 2 shows evidence of two features that bear a slight resemblance to the logos of ICRAR and SHAO. When the data set is simulated and imaged completely the logos will be clearly visible.

Introduction to Imaging and Calibration Pipelines

One aspect of the workshop was to foster collaborations between group members. One element of this collaboration was introducing group members to software packages and data formats. This was achieved via the processing of the datasets above and smaller test datasets, with an aim to widen the usage of these software packages.

This work is ongoing and will enable group members to convert simulation data formats into UVFITS and measurement set (MS) formats for processing via standard pipelines. We will also provide a platform to enable bespoke imaging pipelines to be compared to other packages.

Non-Imaging Analysis

Radio continuum and spectral line surveys will be the predominant data product of the SKA. However time domain analysis is becoming more interesting to the community, especially as the population of transient objects is just starting to be investigated. commensal transient studies will receive correlated visibilities for pulse detection and time-spatial localization. Techniques can be applied to discover both isolated and repeating events. Robust, scalable time domain algorithms including periodogram and wavelet analysis are being tested using statistical methods to identify periodic structures in baseline averaged time series.

In this workshop we applied some of these techniques to data from the SKA precursors including the MWA and ASKAP. The power as a function of frequency channel and reception time for a given baseline was subjected to wavelet thresholding to reduce the noise floor including weak radio frequency interference (RFI) and ensuring that strong RFI can be clearly identified for flagging. Upon stacking these thresholded slices for each baseline into data cubes, it is proposed to identify connected regions and their evolution, useful in identifying RFI and in de-dispersion based strong signal reconstruction for studies including pulsar searches and fast radio burst (FRB) detections. These developments complement imaging pipelines and can serve as a preliminary transients identifier before additional processing with imaging methods owing to their possible near real-time deployment. They can enable a wide range of studies including active galactic nuclei physics, interstellar and intergalactic medium properties, pulsar physics, and identification of FRB progenitors.

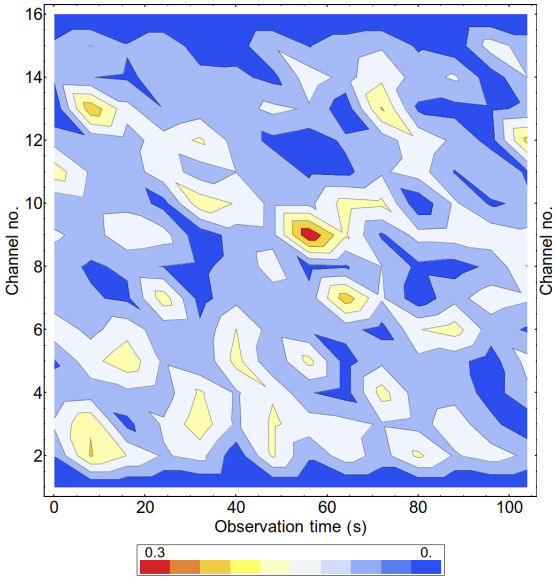

Figure 3: Power on an MWA baseline as a function of time and frequency

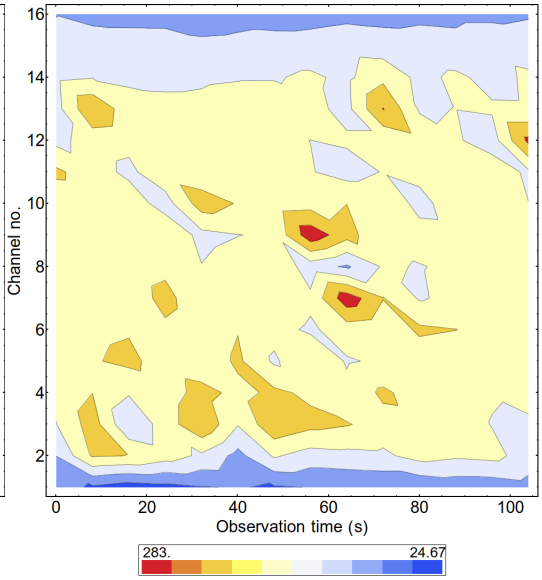

Figure 4: Thresholded power on an MWA baseline. Transient RFI events are clearly visible.

Future Work

Over the next few months we intend to further pursue these collaborations. The SKA data simulation work is already ongoing. We expect to be able to demonstrate ASKAPSoft processing of our collaborators simulation data and comparison between processing pipeline products to also continue on this timescale. The non-imaging analysis will also continue, with some processing of precursor data for transient sources intended.