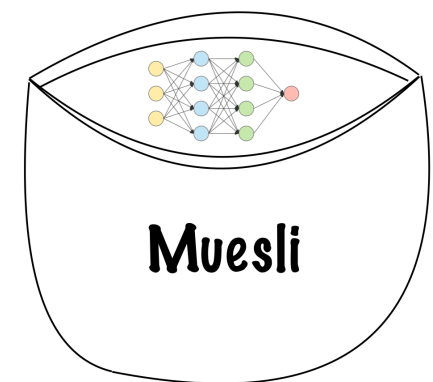
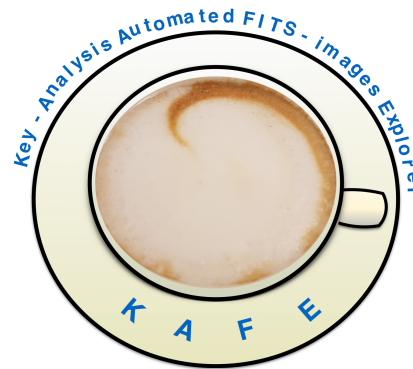


Advanced archival data analysis tools

(The BREAKFAST SUITE)

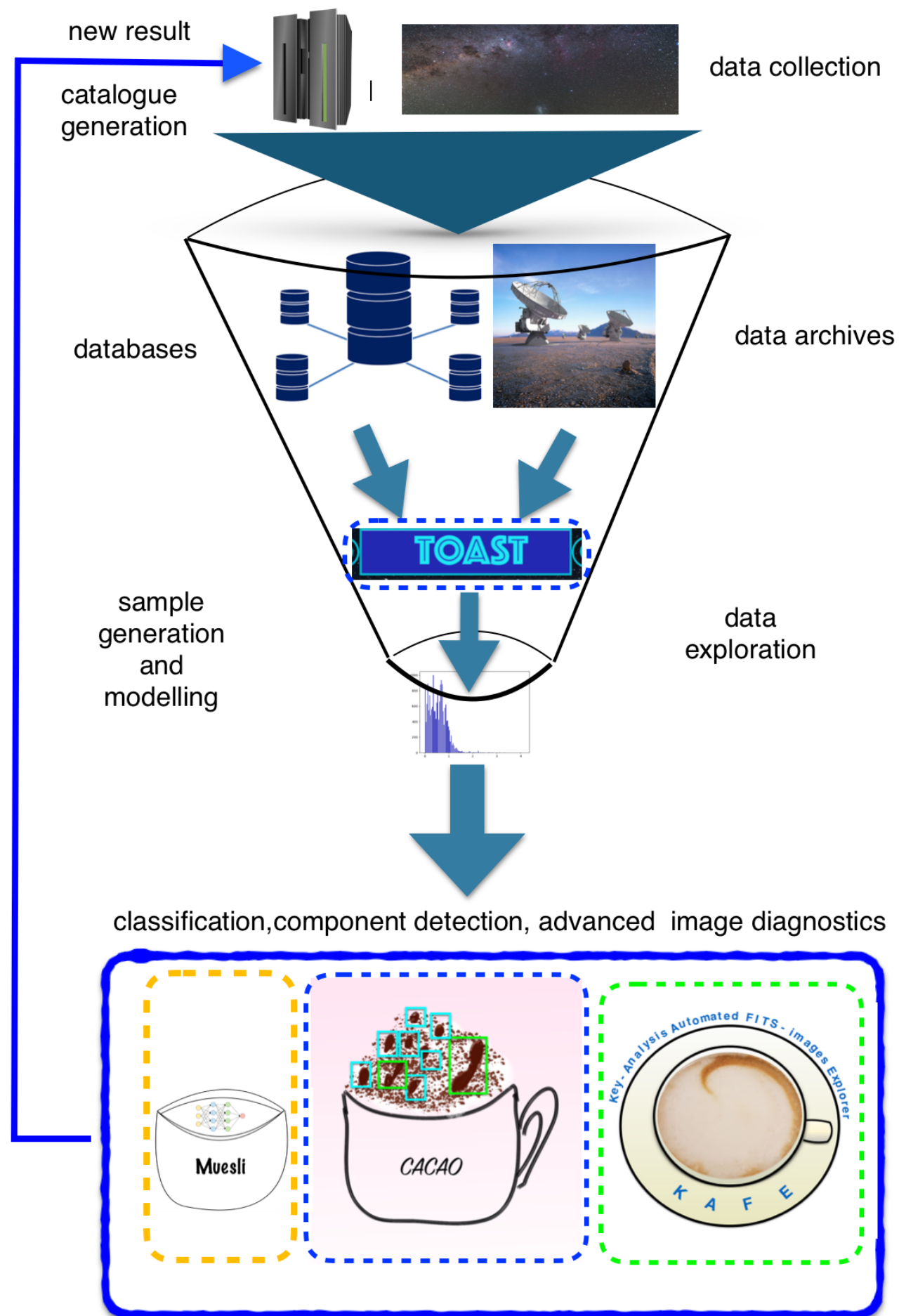


Sandra Burkutean

INAF-Istituto di Radioastronomia, Italian ALMA Regional Center, Bologna

Archival database structure in the Big Data era

- * How do we connect previous results with data archives ?
 - * What kind of data products do we need ?
- * How do we present the data products, given the immense data volume expected in the years to come?



Archival database structure in the Big Data era

- * How do we connect previous results with data archives ?

TOAST: Telescope Observational Archive Sample Tool

The TOAST project

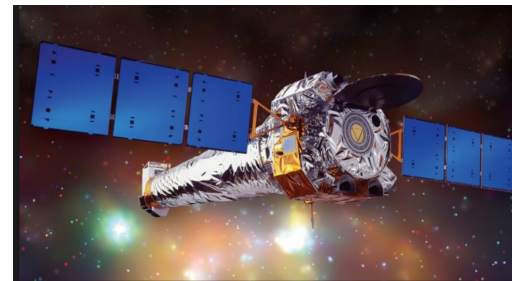
Burkutean in prep.



+



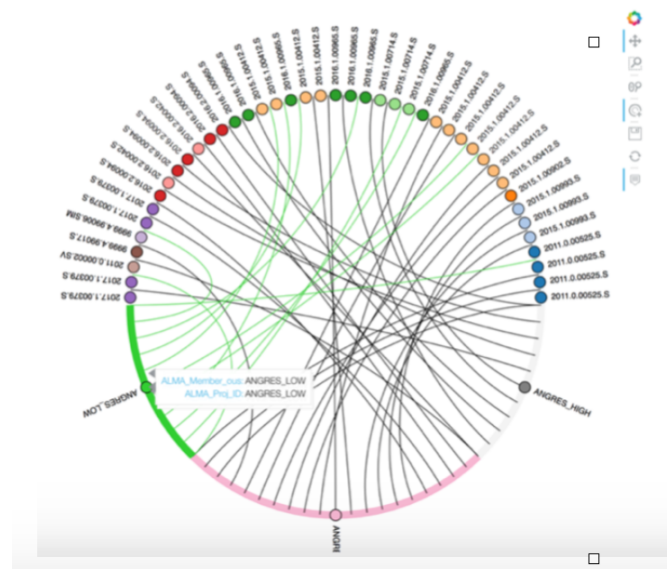
+



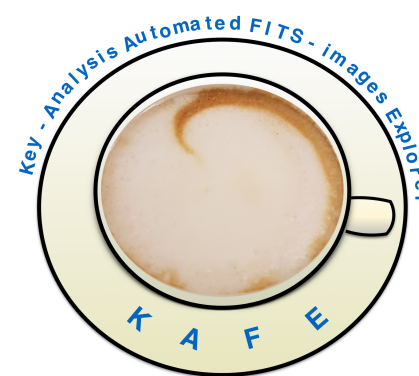
+



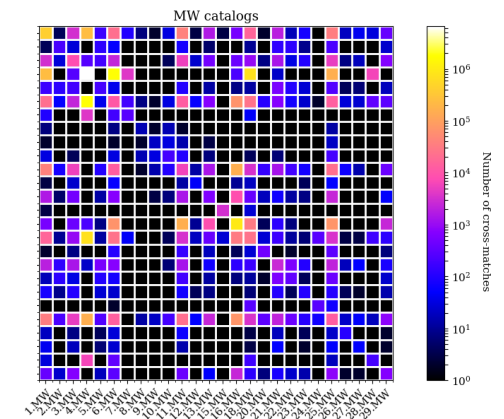
=



+



+



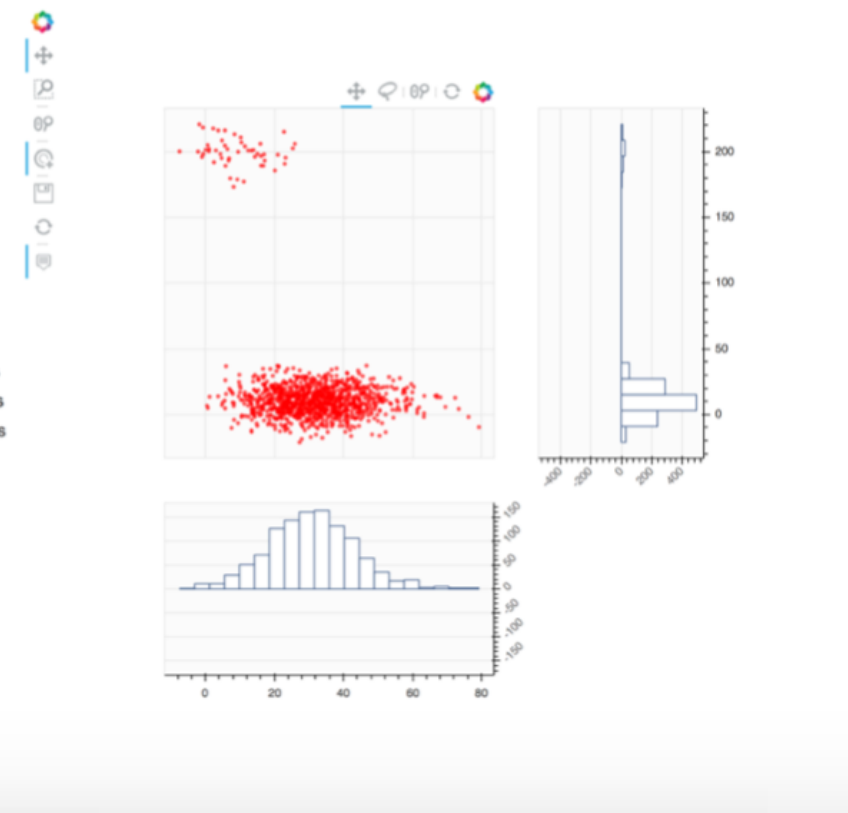
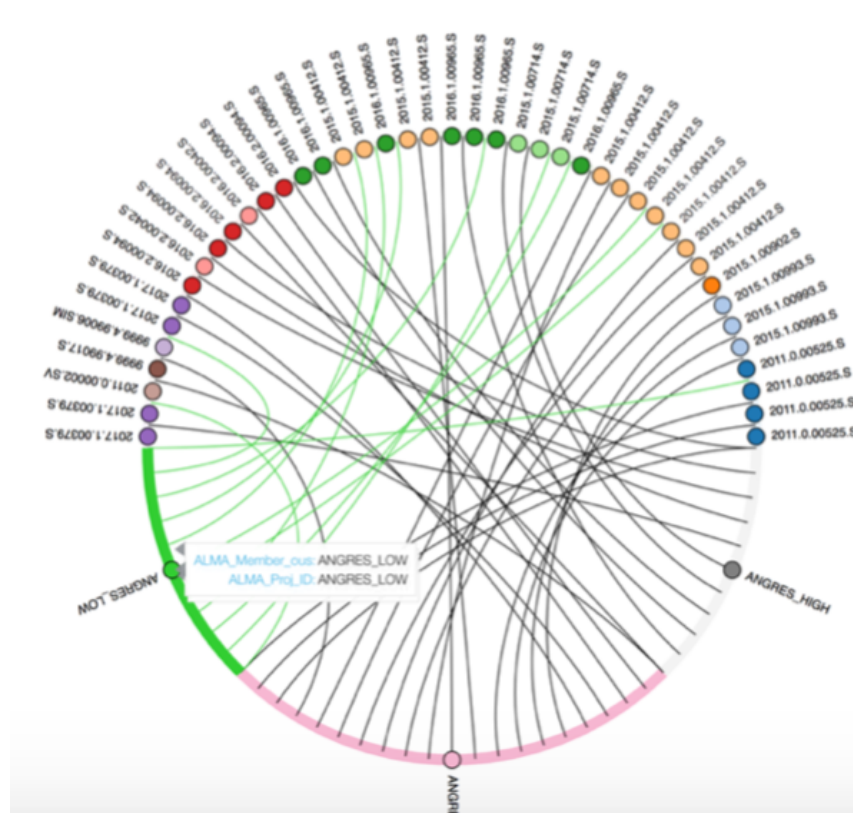
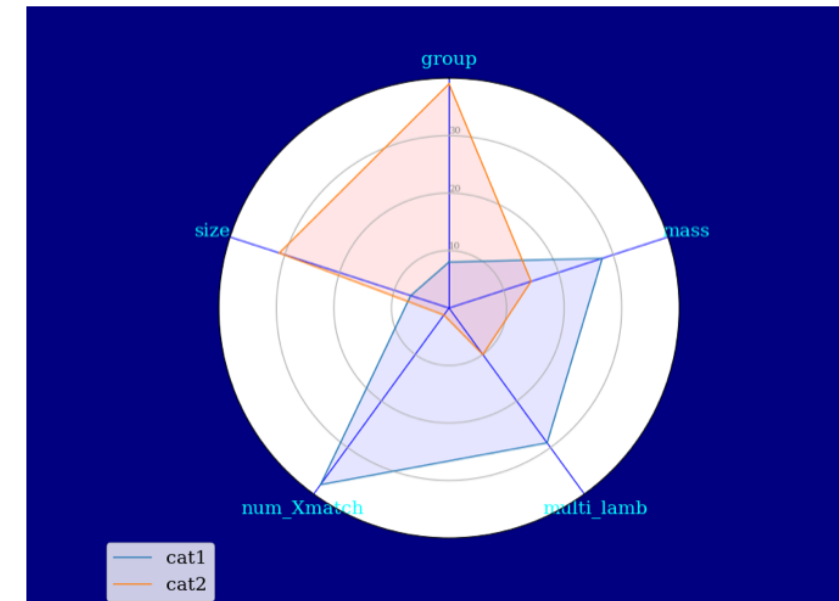
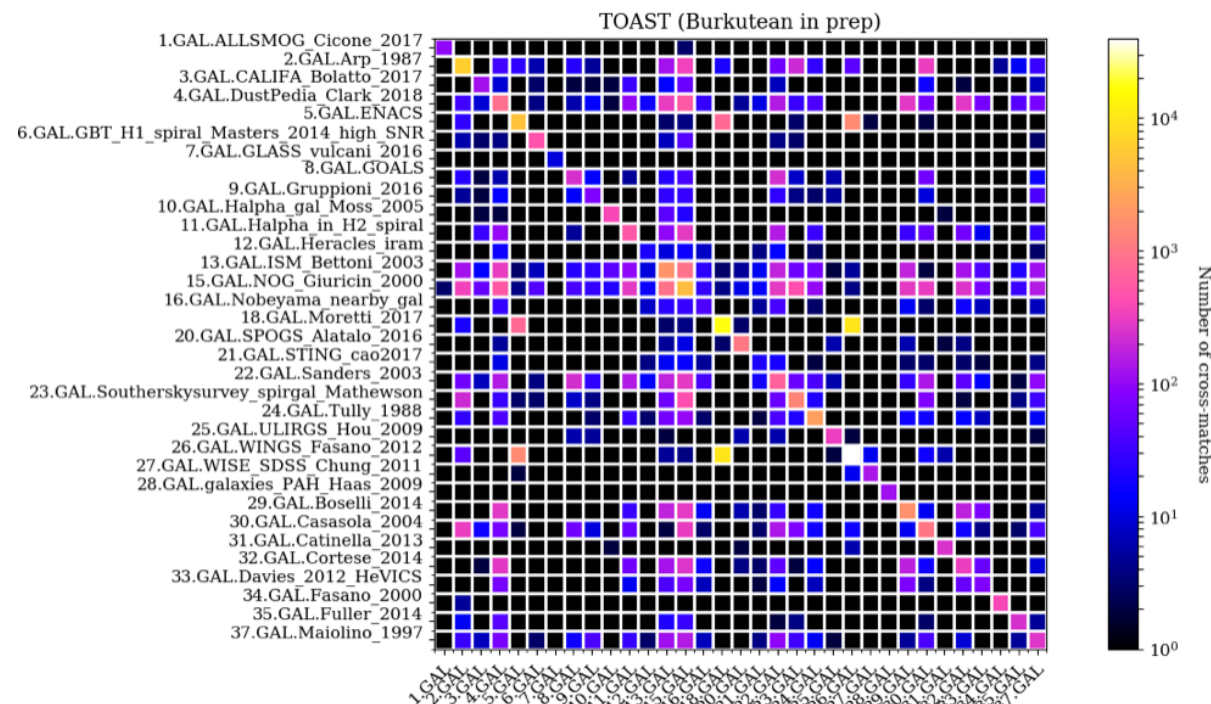
Big DATA
visualisation

Big DATA
analysis

Big DATA
database statistics

TOAST: Telescope Observational Archive Sample Tool

The TOAST project



TOAST: Telescope Observational Archive Sample Tool

The TOAST project

Burkutean in prep.

TOAST



Reset

Continue

TOAST: Telescope Observational Archive Sample Tool

The TOAST project

Burkutean in prep.



<https://www.freepik.com/free-photos-vectors/design>

designed by  freepik.com



Big Data
visualis

Big DATA
base statistics

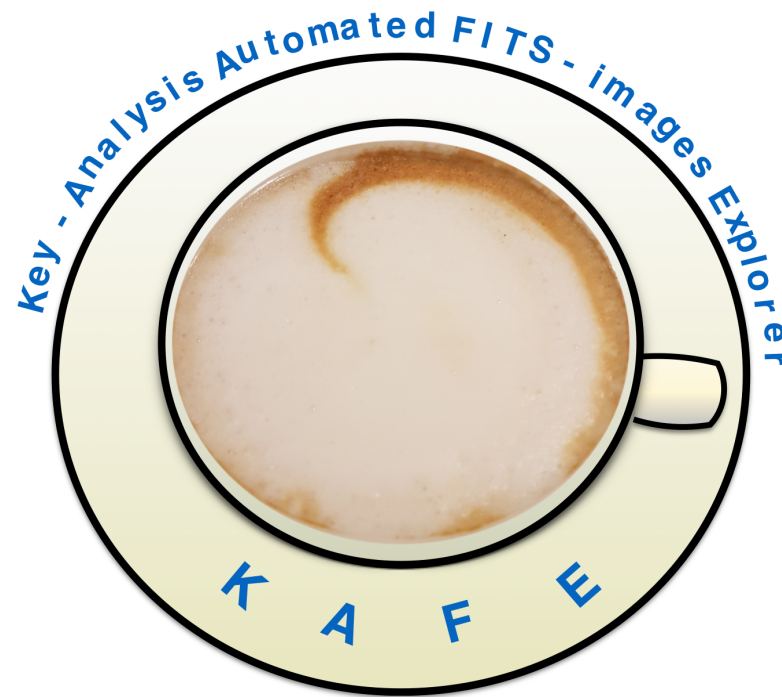
Archival database structure in the Big Data era

- * What kind of data products do we need ?
- * How do we present the data products, given the immense data volume expected in the years to come?

KAFE: automated FITS image analysis + visualisation

The KAFE project

*Burkutean et al., J. Astron.
Telesc. Instrum. Syst. 4(2), 028001 (2018)*



user-generated/archival
FITS-images

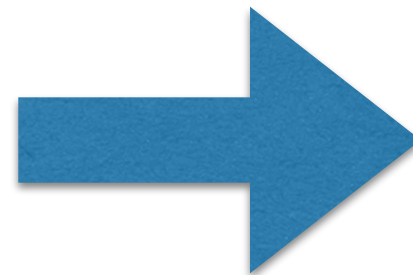
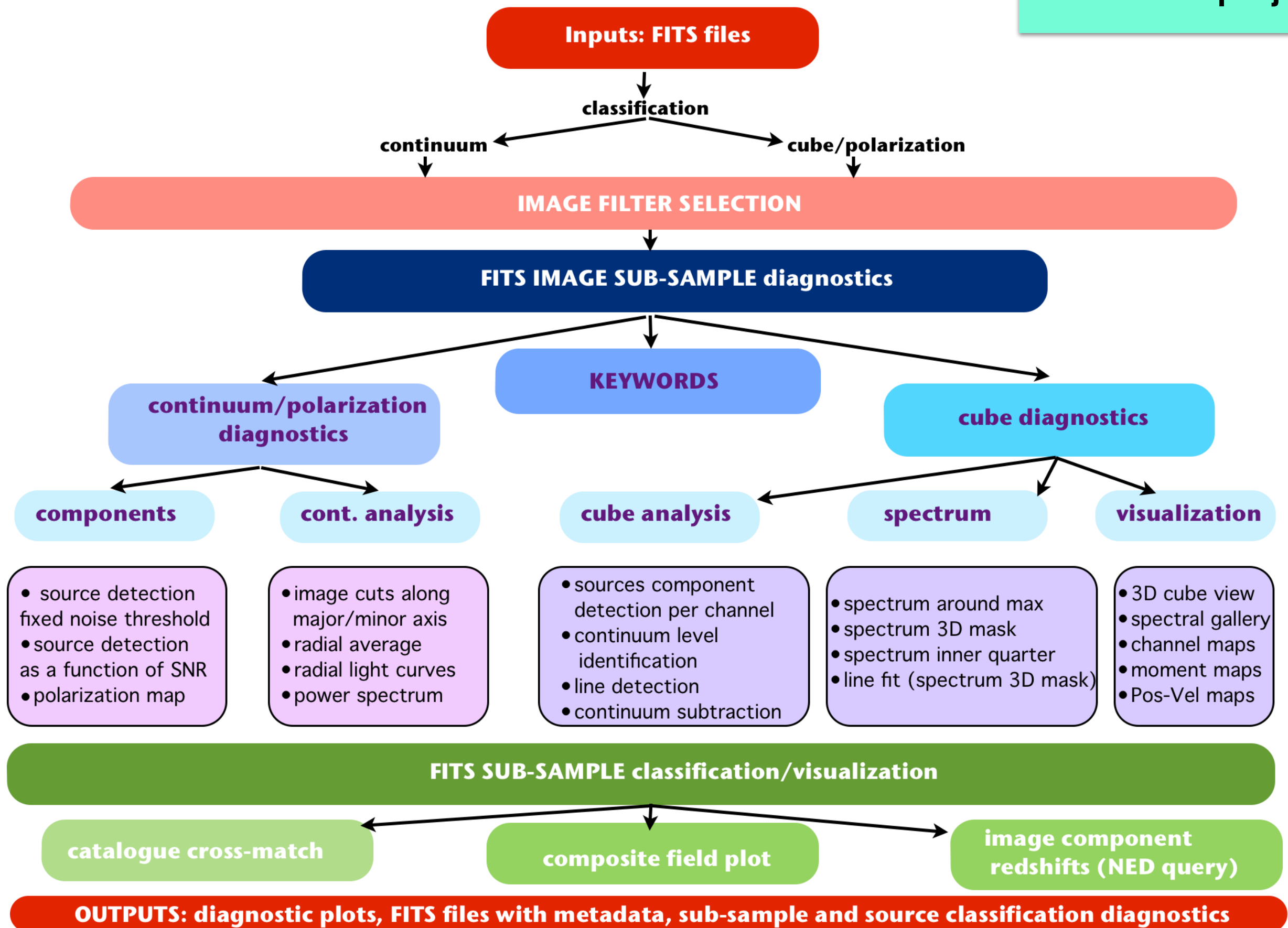


image analysis
for ALMA, JVLA, PdB etc.

“a few clicks”



The KAFE project

send this file:
No file chosen

filters

POS RANGE

ANGRES RANGE

CNTRFREQ RANGE

CHANRMS RANGE

FREQRES RANGE

FLUX TOTAL RANGE

requested keywords

ALL ☐
RA_centre ☐
DEC_centre ☐
SPATRES ☐
BNDCTR ☐
BNDRES ☐
BNDWID ☐
CHANRMS ☐
DYNRANGE ☐
FLUXTOT ☐
DATAMAX ☐
DATAMIN ☐
STOKES ☐

spectrum analysis options

ALL ☐
3D view ☐
continuum subtraction ☐
Channel gallery ☐
Spectrum_3D_mask ☐
Spectrum inner quarter ☐
Spectrum around max ☐
Spectral gallery ☐
3D posvel ☐
moments ☐
PosVel along maj/min axis ☐
Spectral fit ☐
Cube morph ☐

further analysis options

ALL (except LC,3colour) ☐
Source detection ☐
Source detection SNR layer ☐
radial average ☐
Image cuts ☐
power spectrum ☐
Polarization maps ☐
Light curve ☐
3-colour image ☐

catalog selection

HDF ☐
Chandra DFN ☐
FRICAT ☐

ATHDFSOID ☐
COSMOS Chandra bright src ☐
FRIICAT ☐

HUDF ☐
COSMOS VLA deep ☐
BzCAT ☐

Chandra DFS ☐
FERMILAC ☐
SPTSZSPSC ☐

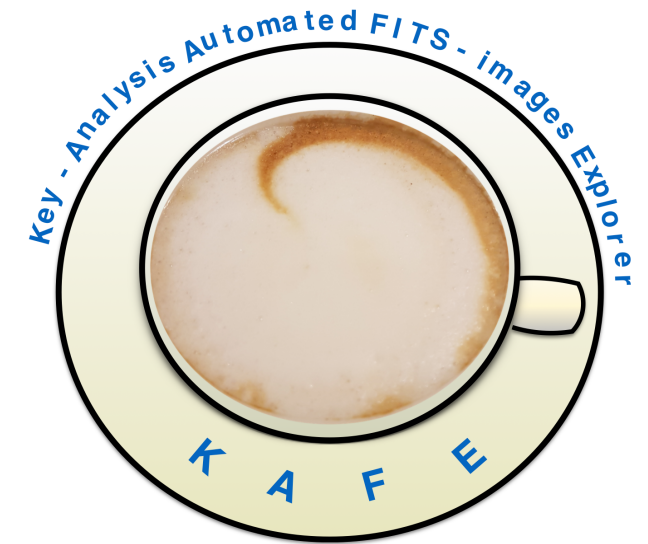
cross-match query and output specifications

catalogue cross-match ☐
RA-redshift pie plot ☐

NED redshift catalogue cross-match ☐
Mollweide all-sky sample plot ☐

NED photometry cross-match ☐
insert KAFE keywords into FITS header ☐

composite field (FOV) plot ☐
*.png *.bt *.dat products only ☐

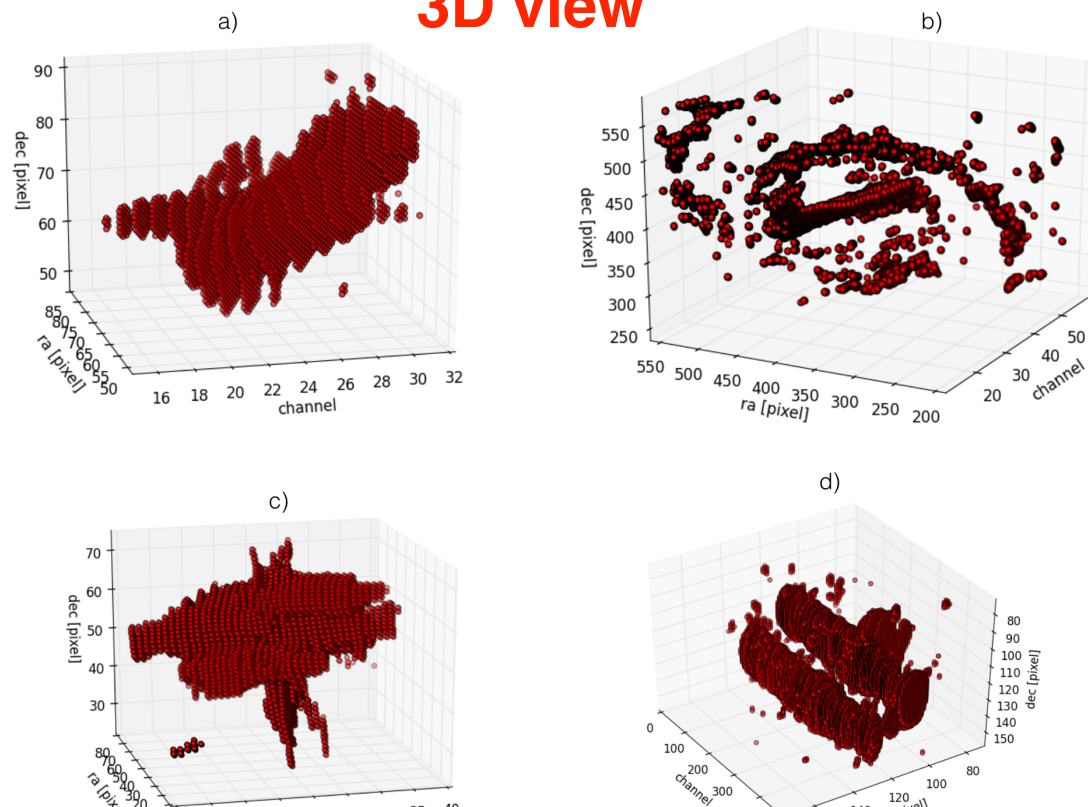


AIMS:

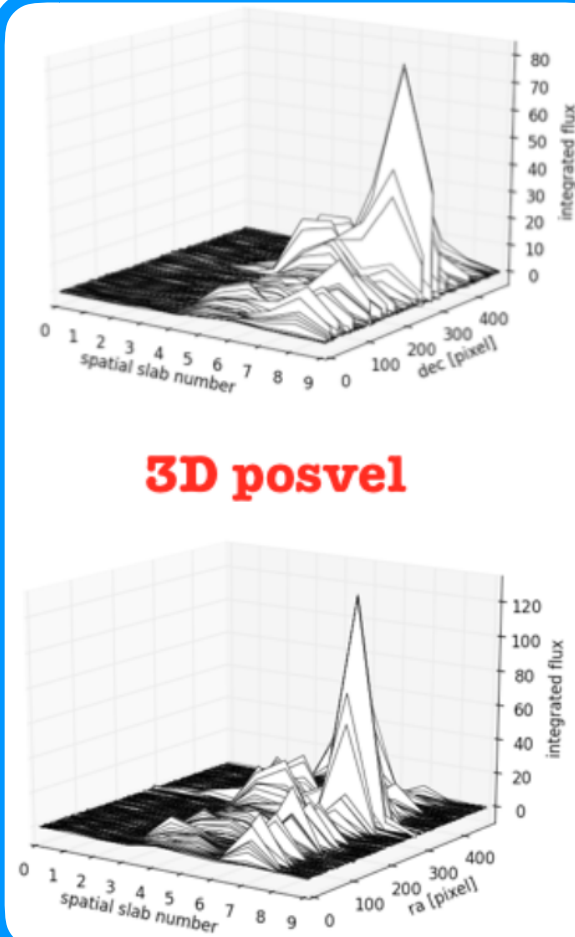
- provide advanced image analysis diagnostic plots in the spatial, spectral and temporal domain for user input FITS images
- provide catalogue cross-matching
- minimal user input required (just tick the boxes) - the image computations and the required parameter settings are fully automated

The KAFE project

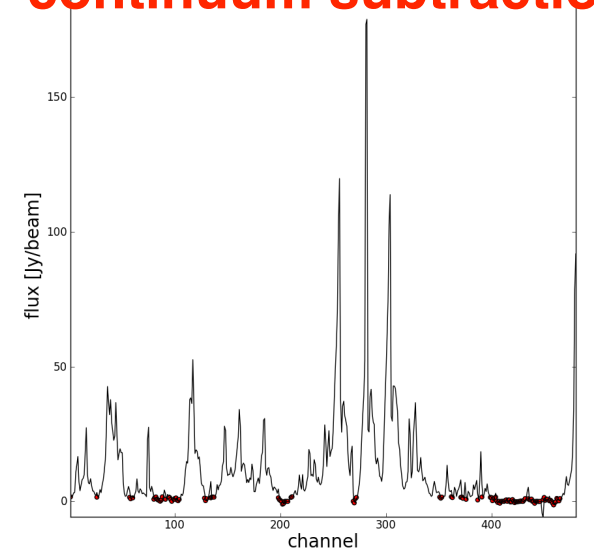
3D view



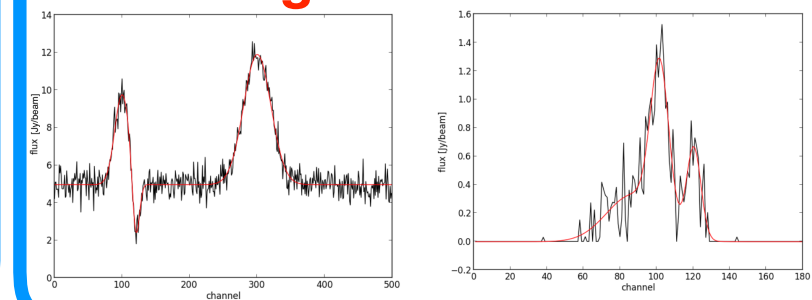
3D posvel



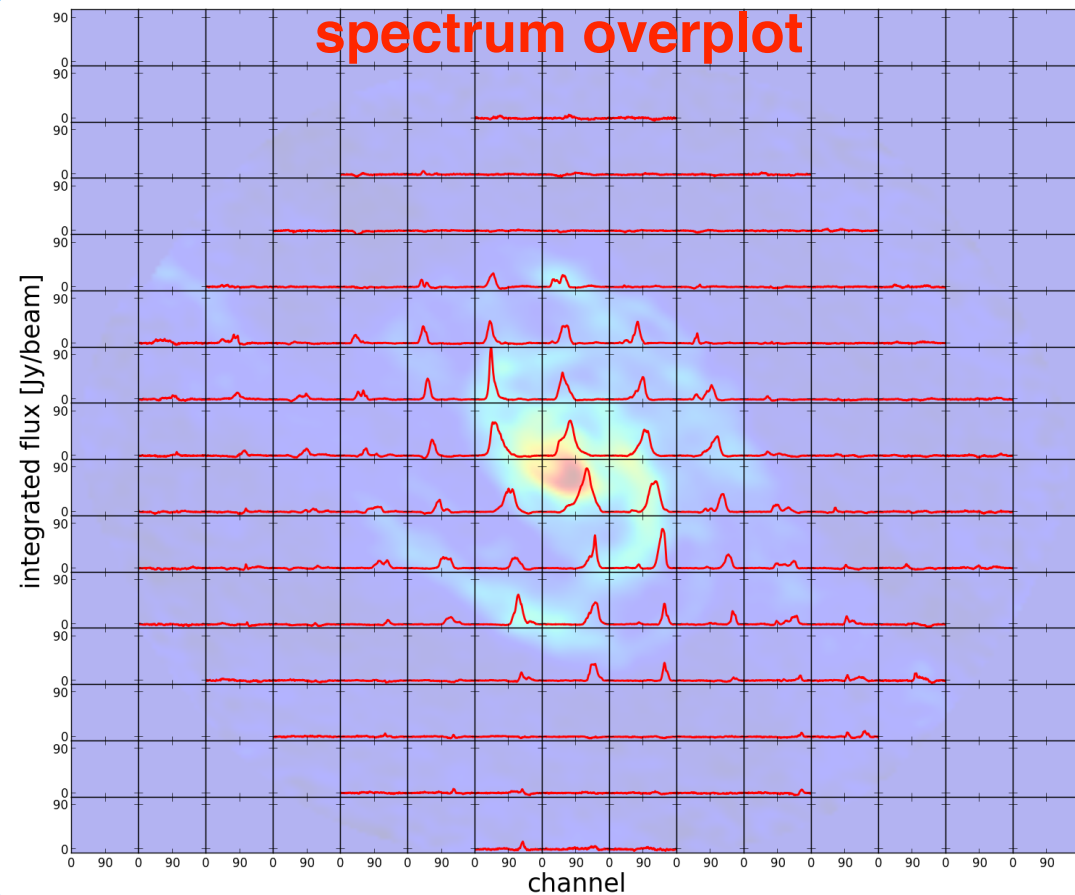
continuum subtraction



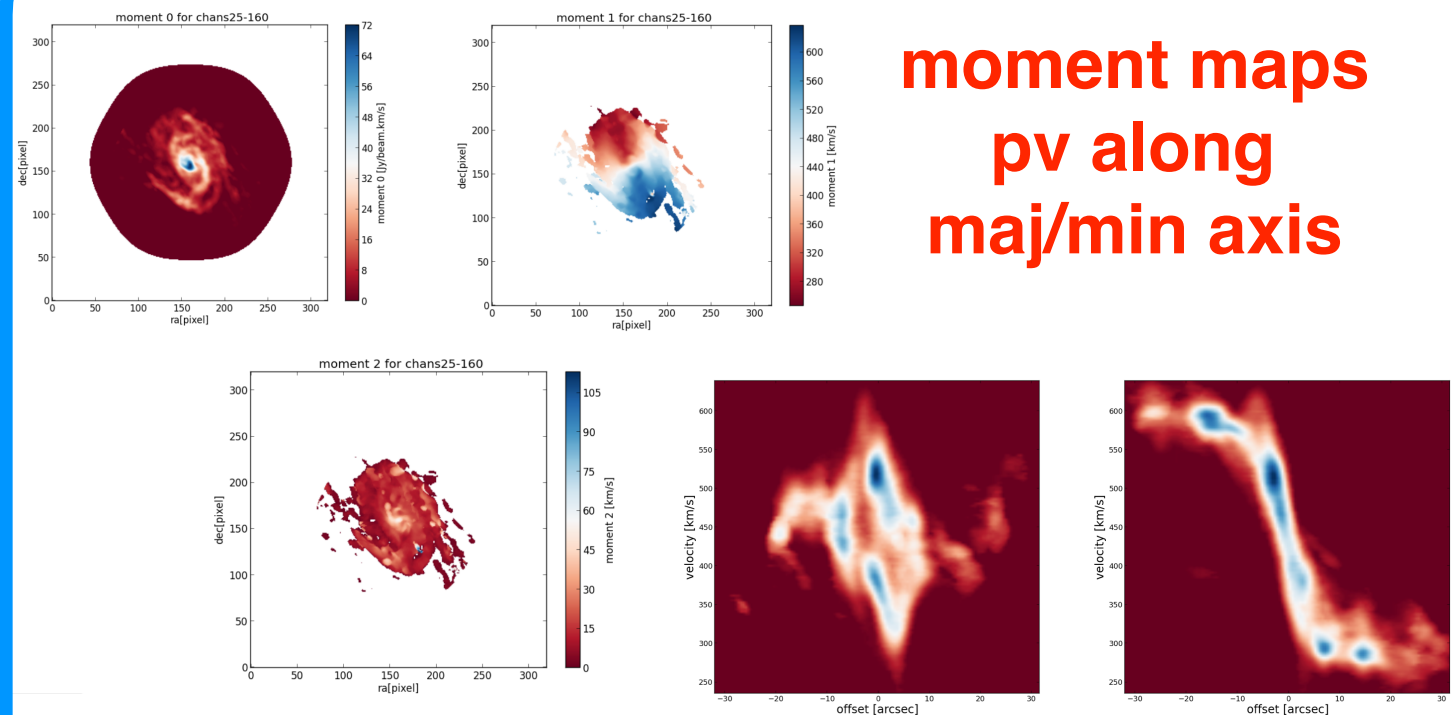
line fitting



spectrum overplot

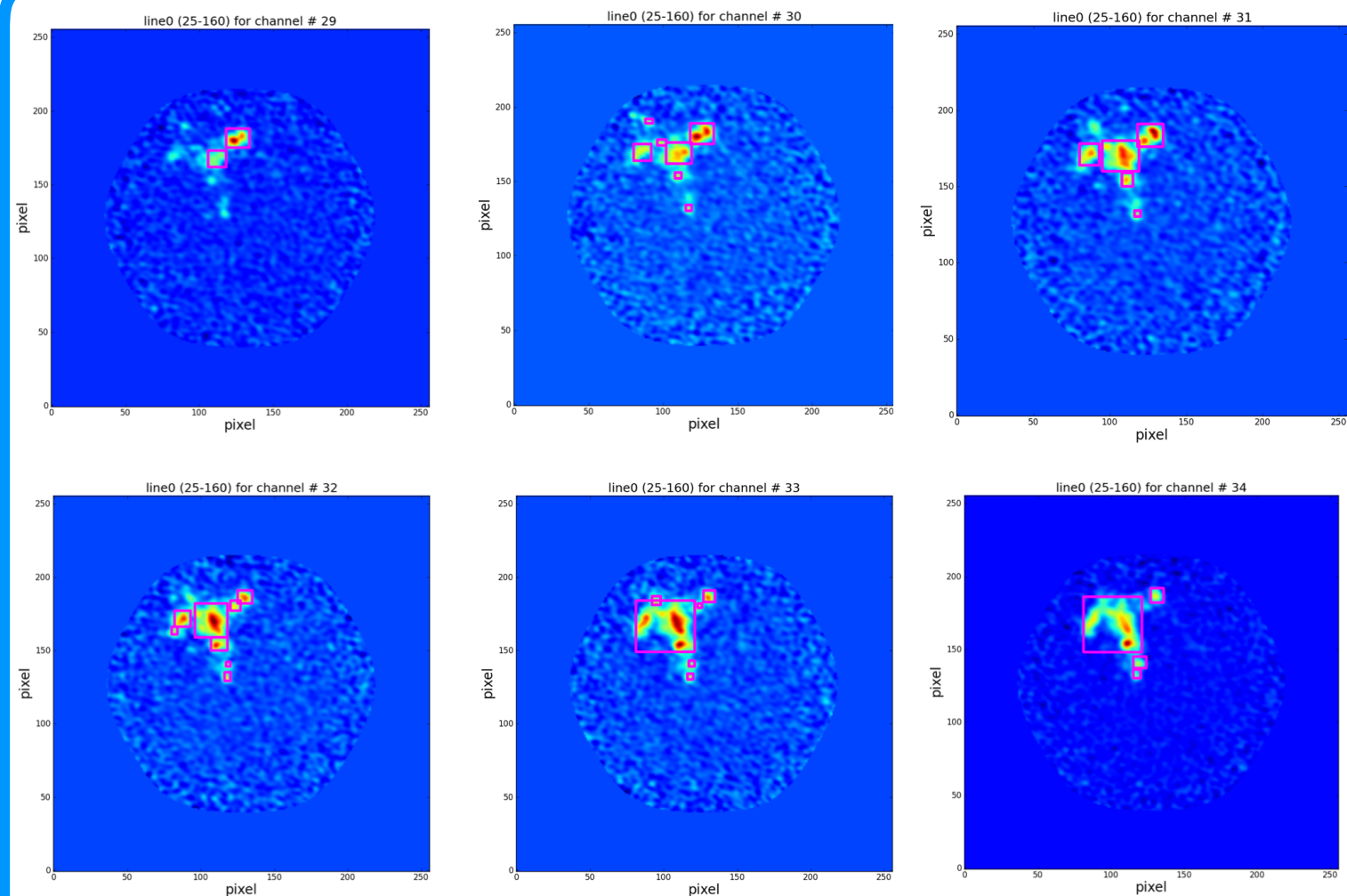


moment maps pv along maj/min axis



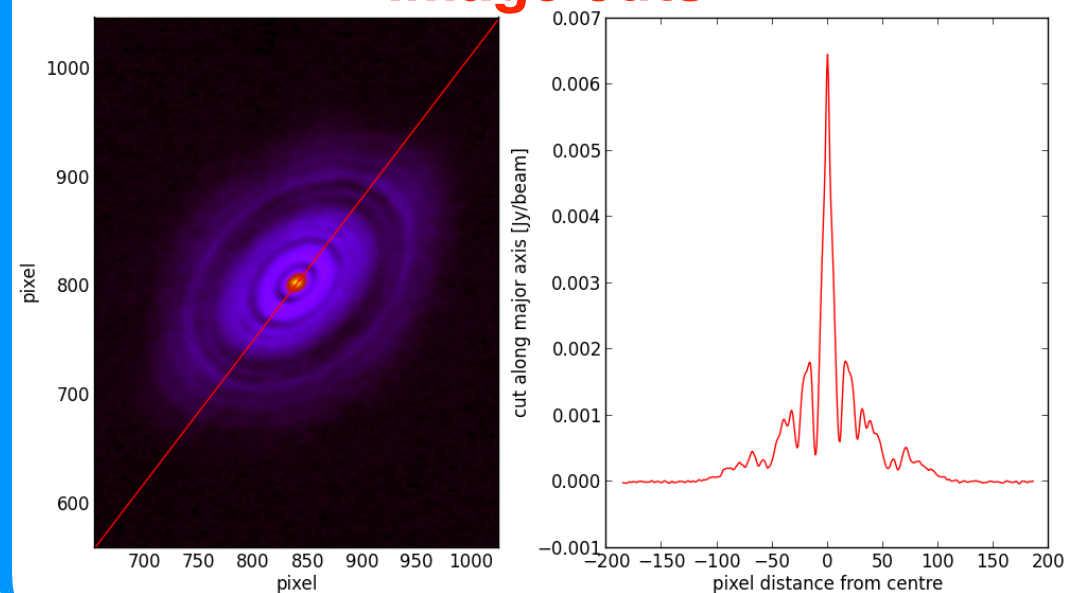
further analysis options

- ALL (except LC, 3colour) ☐
- Source detection ☐
- Source detection SNR layer ☐
- radial average ☐
- Image cuts ☐
- power spectrum ☐
- Polarization maps ☐
- Light curve ☐
- 3-colour image ☐

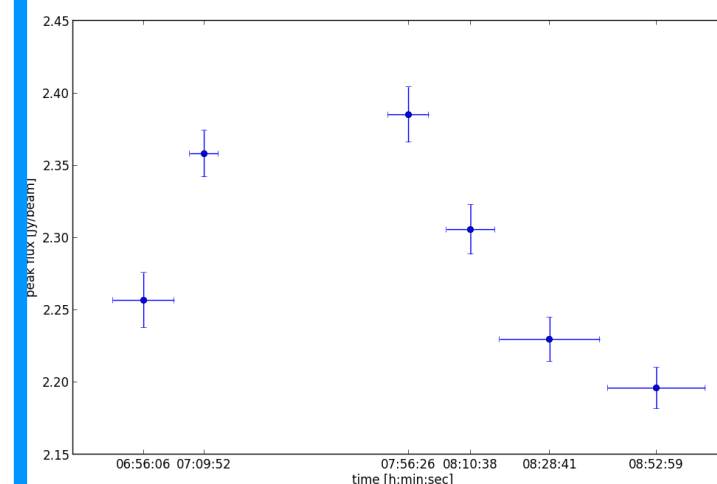


component detection

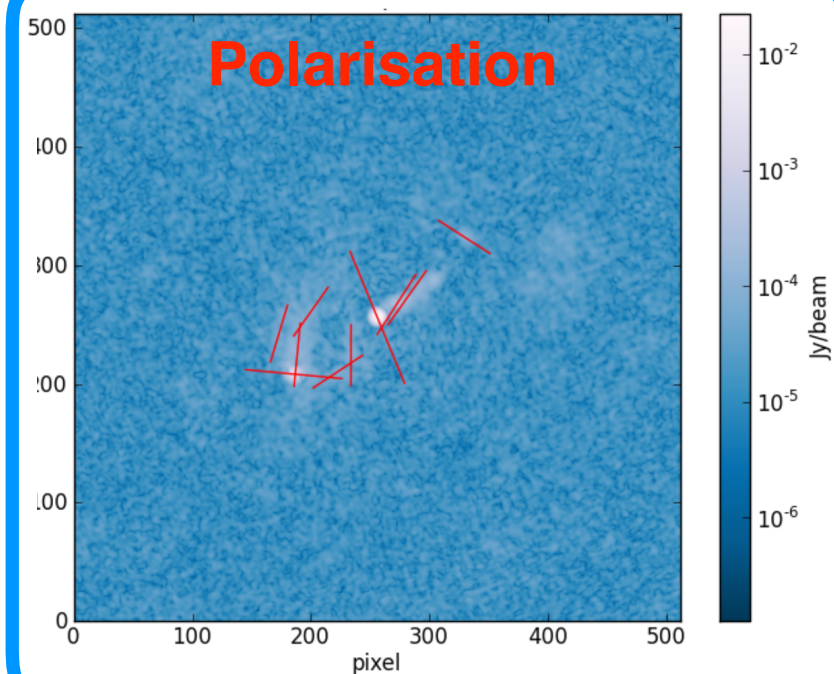
image cuts

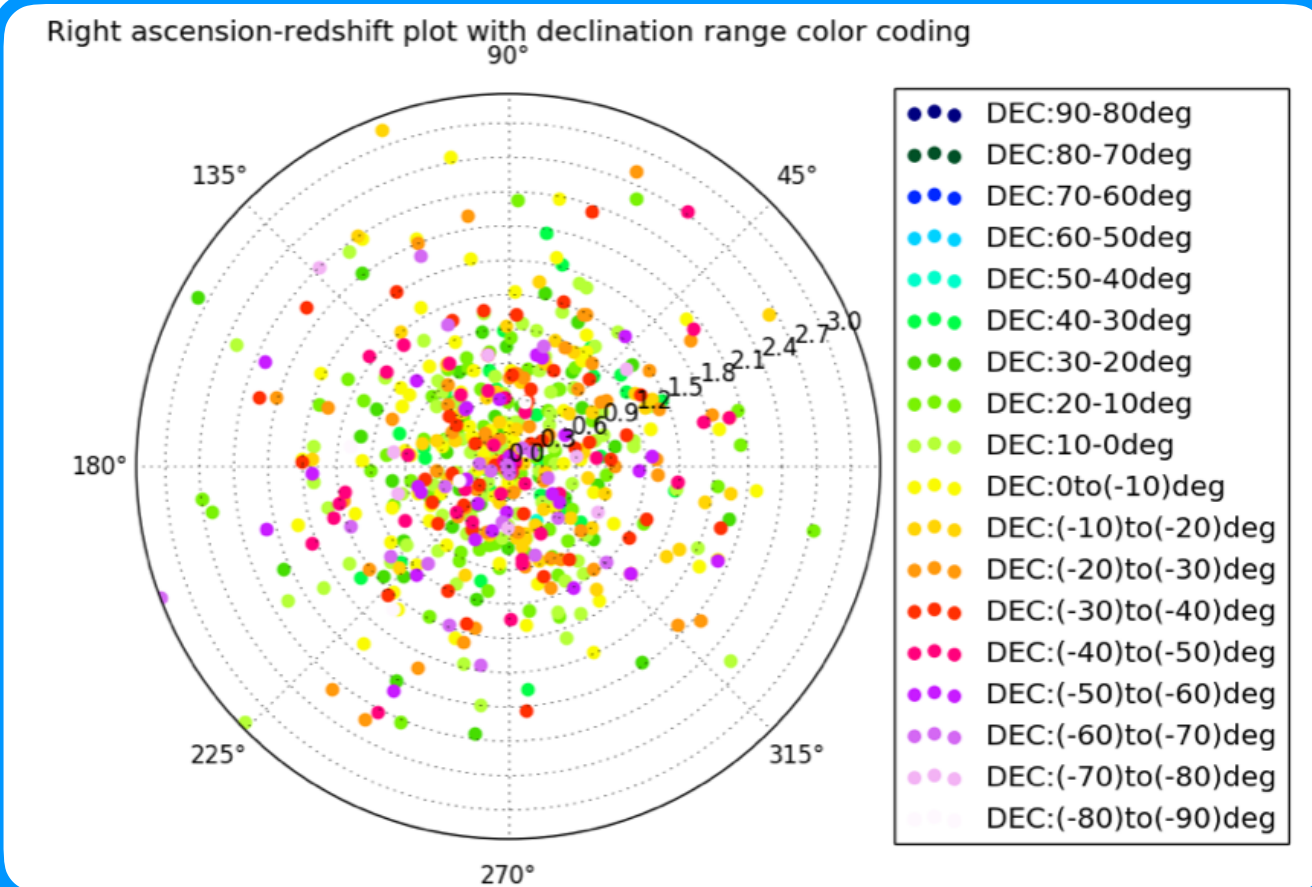
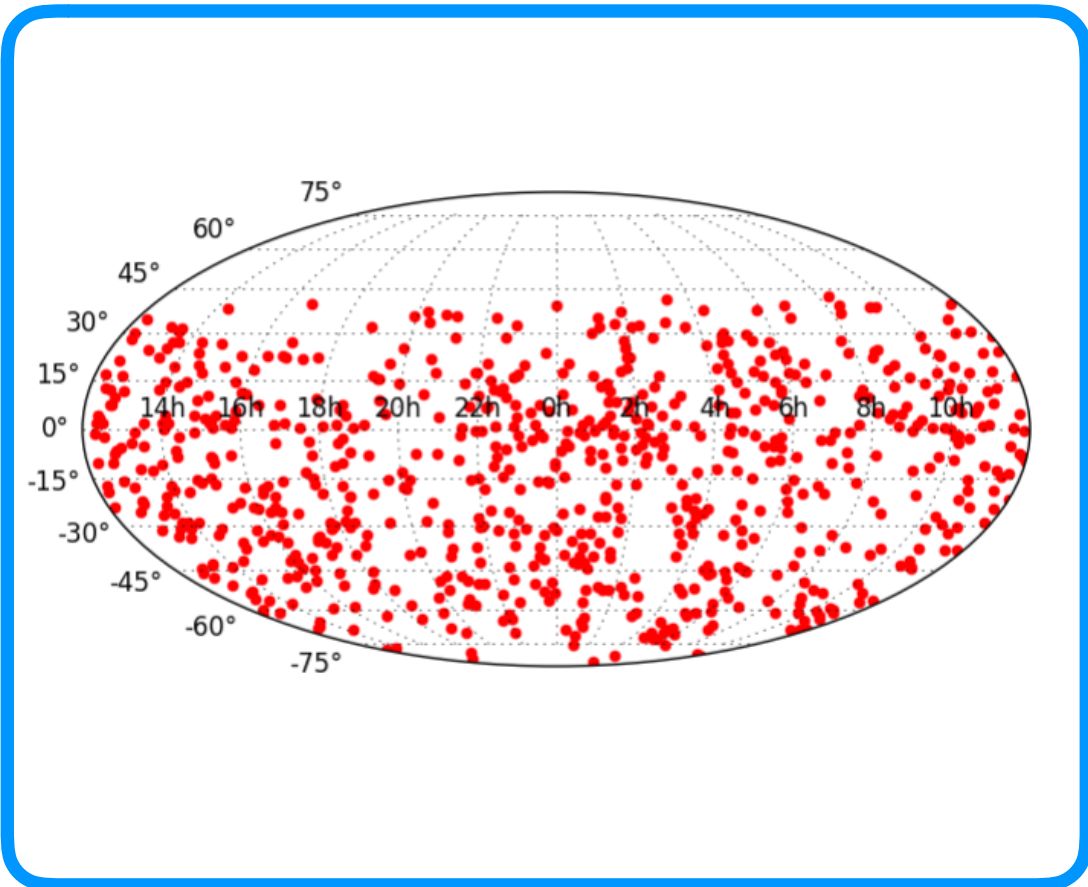
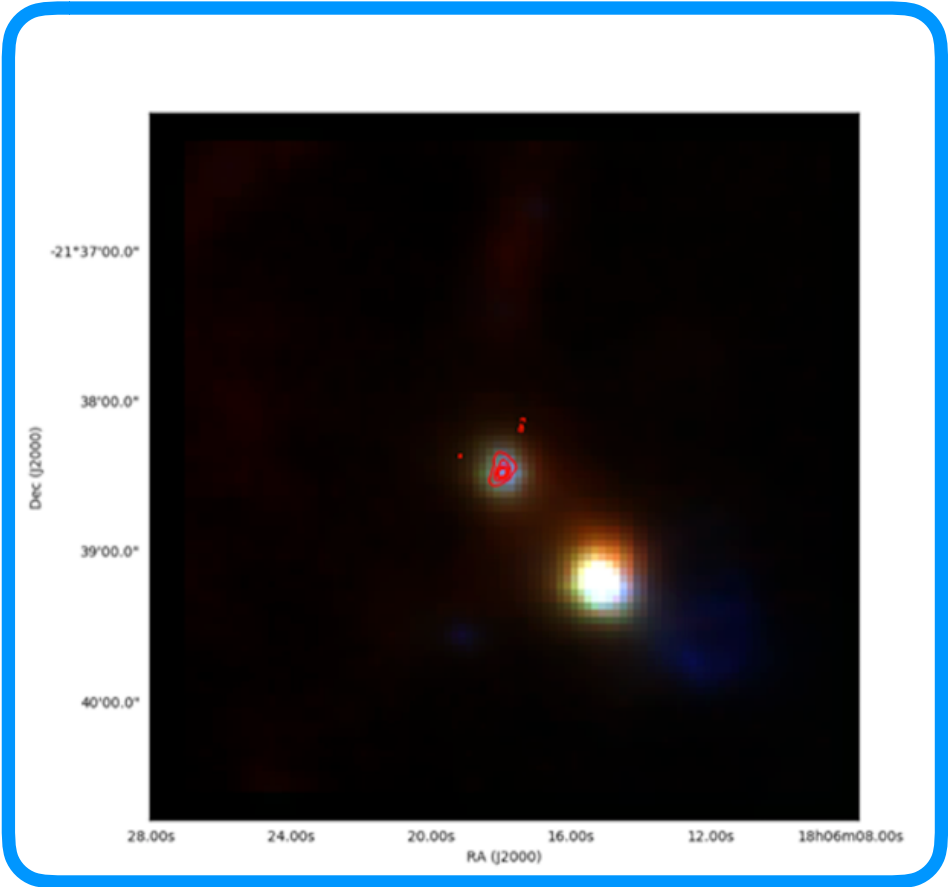
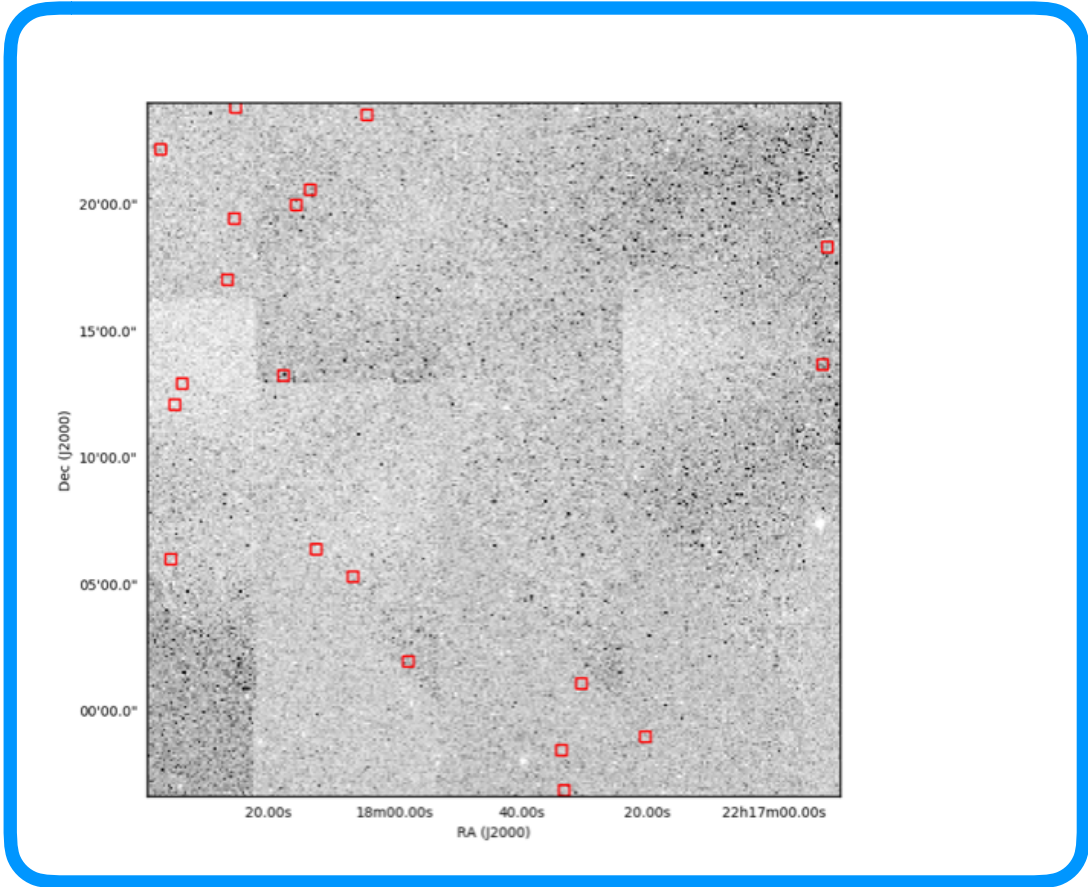


Light curves



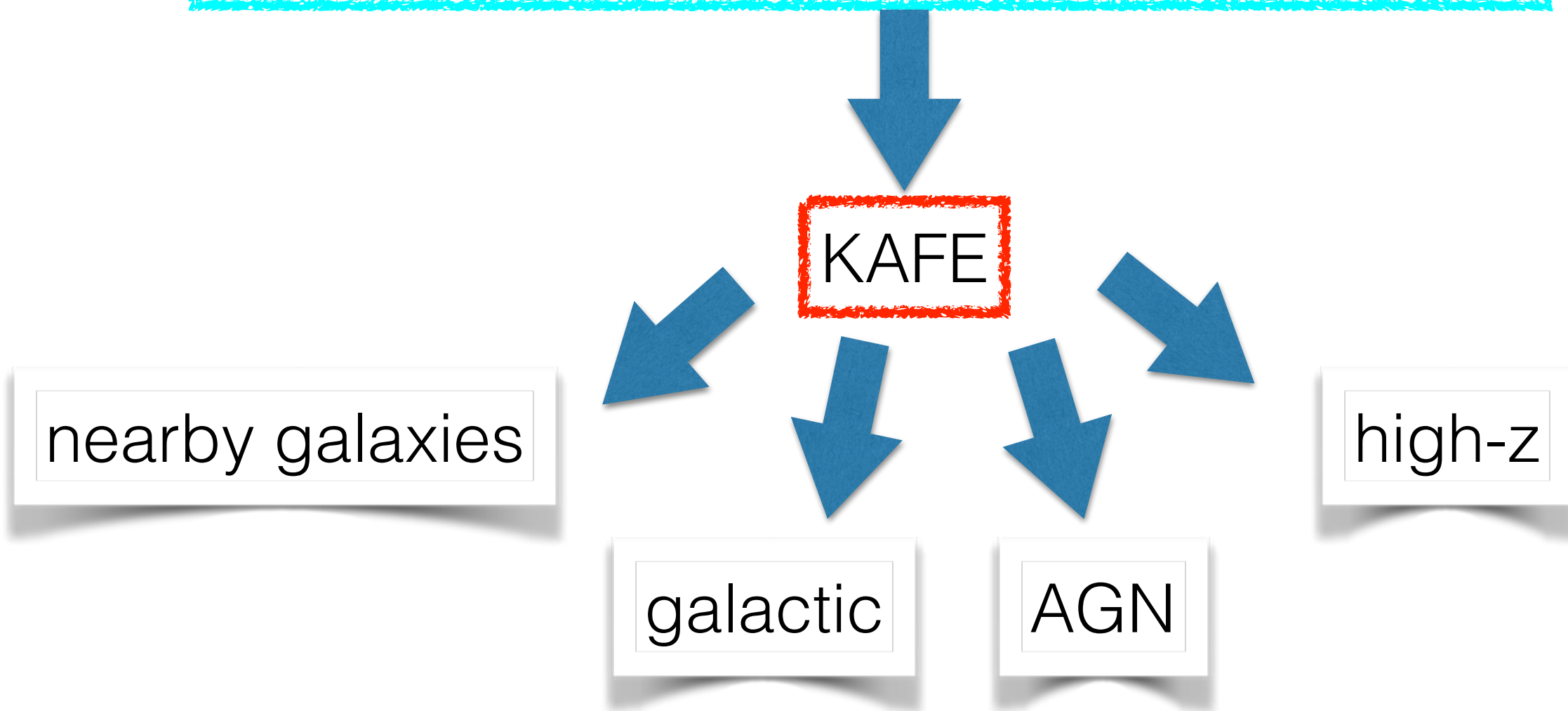
Polarisation





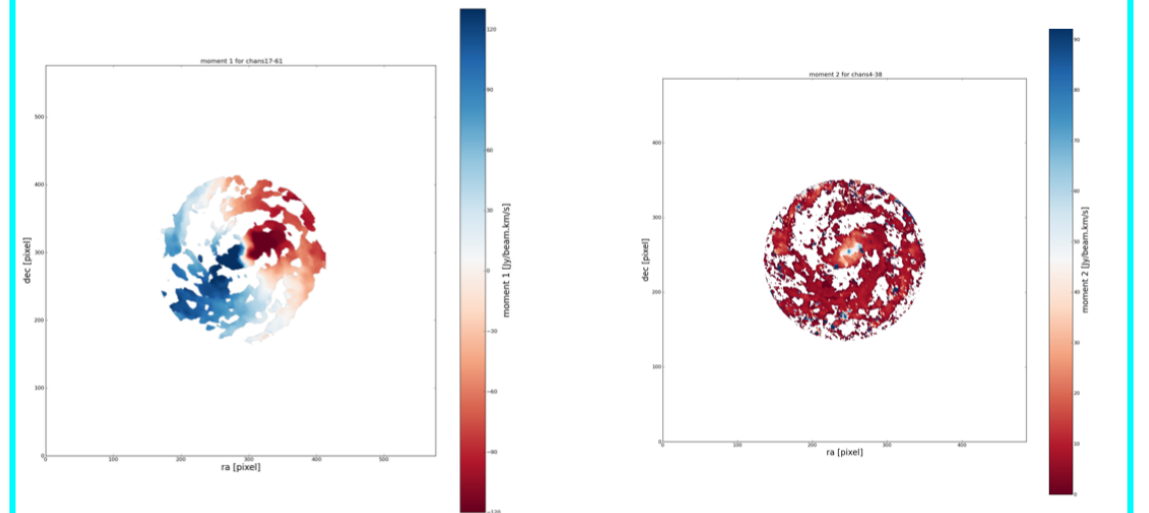
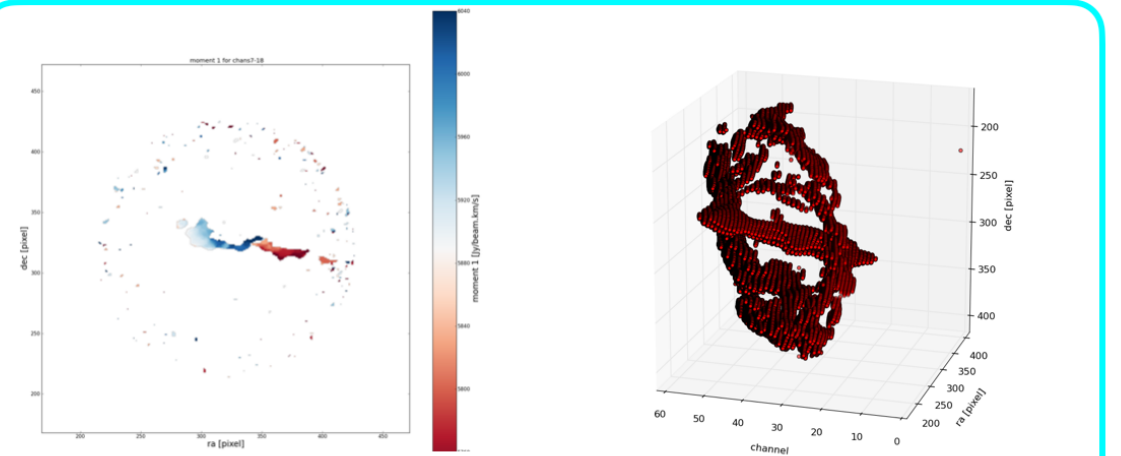
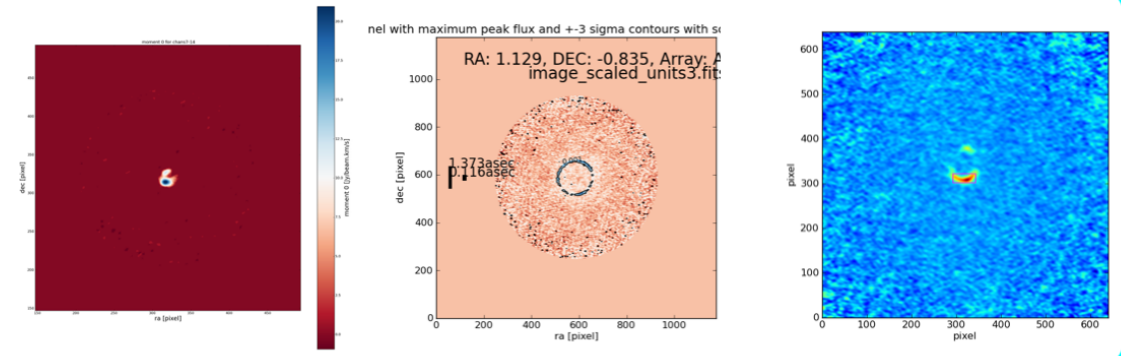
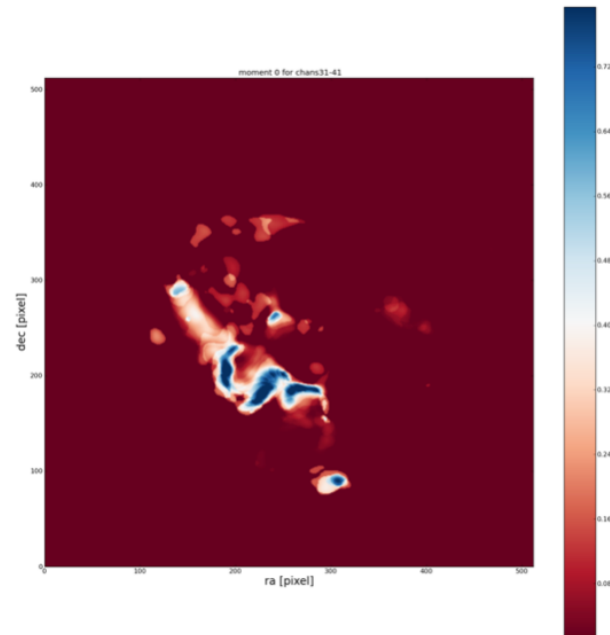
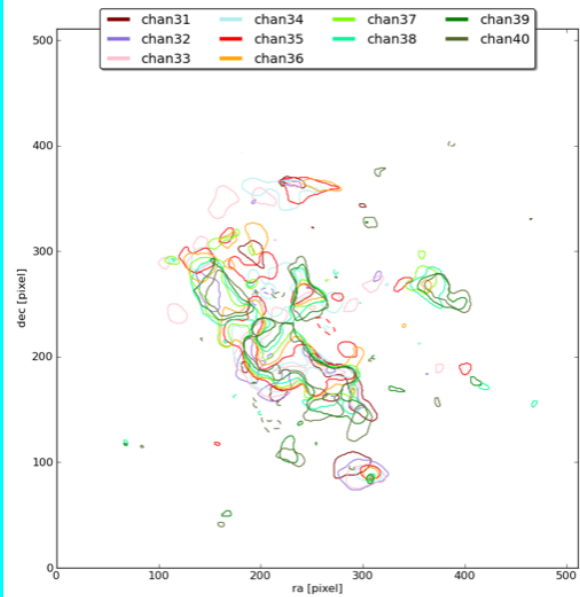
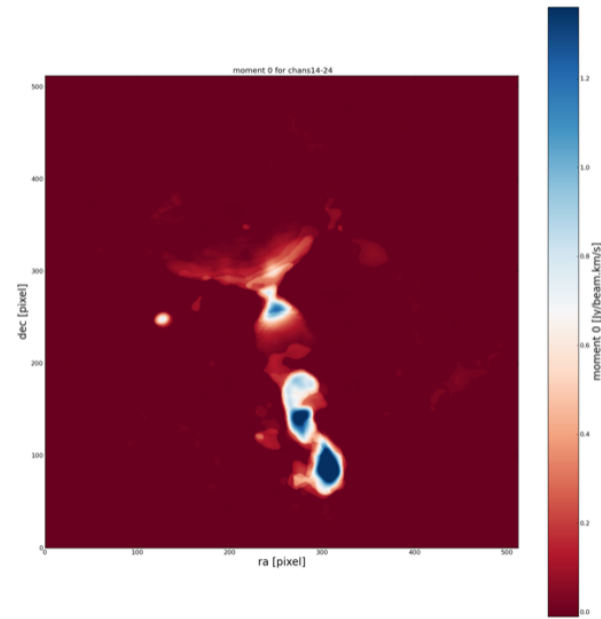
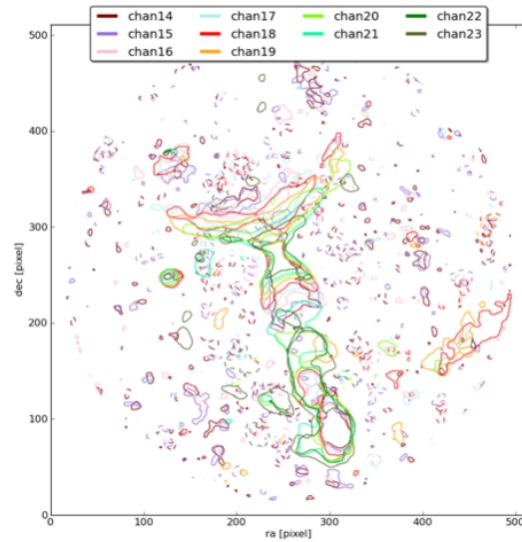
KAFEwithALMA project

Public archival images in the ALMA archive



on Italian ARC cluster + HPC computing time at CHIPP (Trieste)

KAFEwithALMA project



send this file: No file chosen

filters

POS RANGE0

CNTRFREQ RANGE0

FREQRES RANGE0

ANGRES RANGE0

CHANRMS RANGE0

FLUX TOTAL RANGE0

requested keywords

ALL

RA_centre

DEC_centre

SPATRES

BNDCTR

BNDRES

BNDWID

CHANRMS

DYNRANGE

FLUXTOT

DATAMAX

DATAMIN

STOKES

spectrum analysis options

ALL

3D view

continuum subtraction

Channel gallery

Spectrum_3D_mask

Spectrum inner quarter

Spectrum around max

Spectral gallery

3D posvel

moments

PosVel along maj/min axis

Spectral fit

Cube morph

further analysis options

ALL (except LC,3colour)

Source detection

Source detection SNR layer

radial average

Image cuts

power spectrum

Polarization maps

Light curve

3-colour image

catalog selection

HDF

Chandra DFN

FRICAT

ATHDFSOID

COSMOS Chandra bright src

FRIICAT

HUDF

COSMOS VLA deep

BzCAT

Chandra DFS

FERMILAC

SPTSZSPSC

cross-match query and output specifications

catalogue cross-match

RA-redshift pie plot

NED redshift catalogue cross-match

Mollweide all-sky sample plot

NED photometry cross-match

insert KAFE keywords into FITS header

composite field (FOV) plot

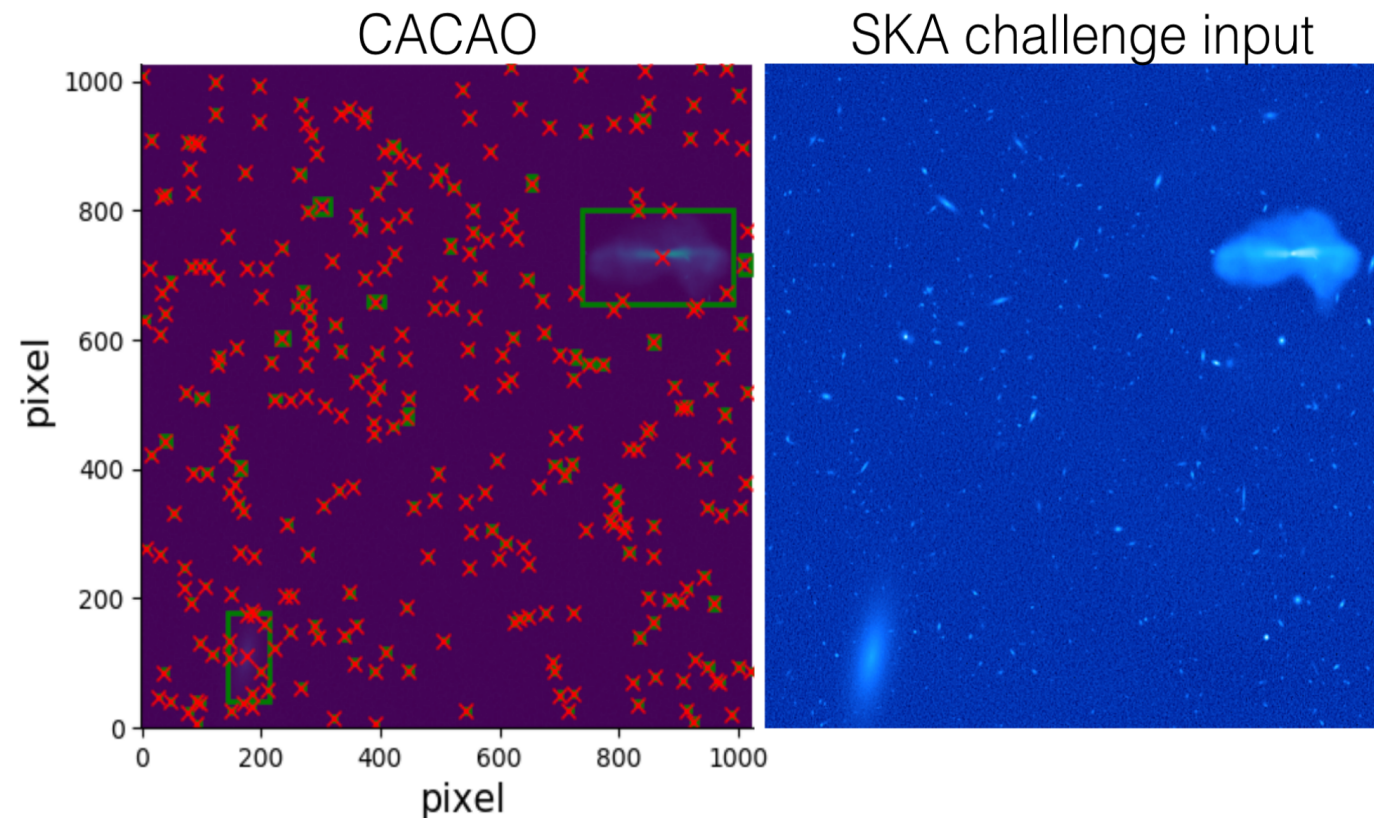
*.png *.txt *.dat products only

process



Please write to kafe@ira.inaf.it for access information to the web interface.

CACAO: The Complete Automated Classification of Astronomical Objects Tool



developed for the 1st SKA Challenge

written from scratch and fully parallelized

object detection, description and classification

further developments and applications to real data ongoing

machine learning application with MUESLI in planning

CACAO: The Complete Automated Classification of Astronomical Objects Tool

INPUT image

Data preparation

pbcor and metadata checks

```

NAXIS1 = 1024 / number of array dimensions
NAXIS2 = 1024
WCSAXES = 2 / Number of coordinate axes
CRPIX1 = 511.75 / Pixel coordinate of reference point
CRPIX2 = 511.75 / Pixel coordinate of reference point
CDEL1 = -0.000167847 / [deg] Coordinate increment at reference point
    
```

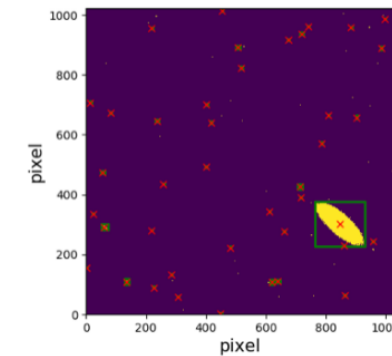
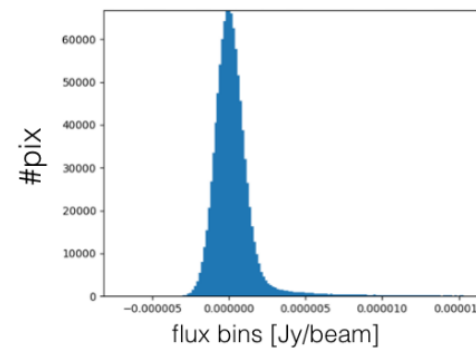
```

CDEL2 = 0.000167847 / [deg] Coordinate increment at reference point
CUNIT1 = 'deg' / Units of coordinate increment
CUNIT2 = 'deg' / Units of coordinate increment
CTYPE1 = 'RA---SIN' / Right ascension, orthographic projection
CTYPE2 = 'DEC--SIN' / Declination, orthographic projection
CRVAL1 = 359.66386291092 / [deg] Coordinate value at reference point
    
```

BLOCK 1

(noise diagnostics, source detection)

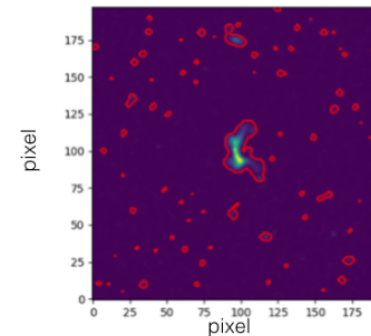
thresholding with examination of skewness and kurtosis
taking into account primary beam correction effects
isolated islet identification with a
diagonalized structure element template



BLOCK 2

(source description)

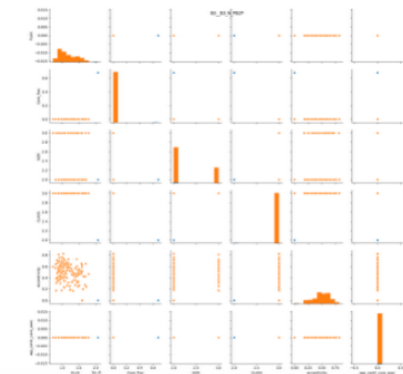
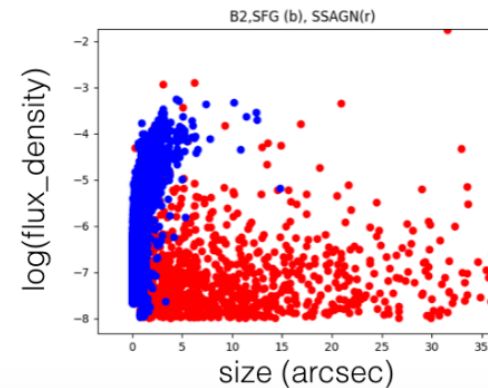
largest angular scale determination along major
and minor axis of emission regions
more sophisticated methods (active contours etc.) are
currently being developed



+ further
future code
developments

BLOCK 3
(source classification)

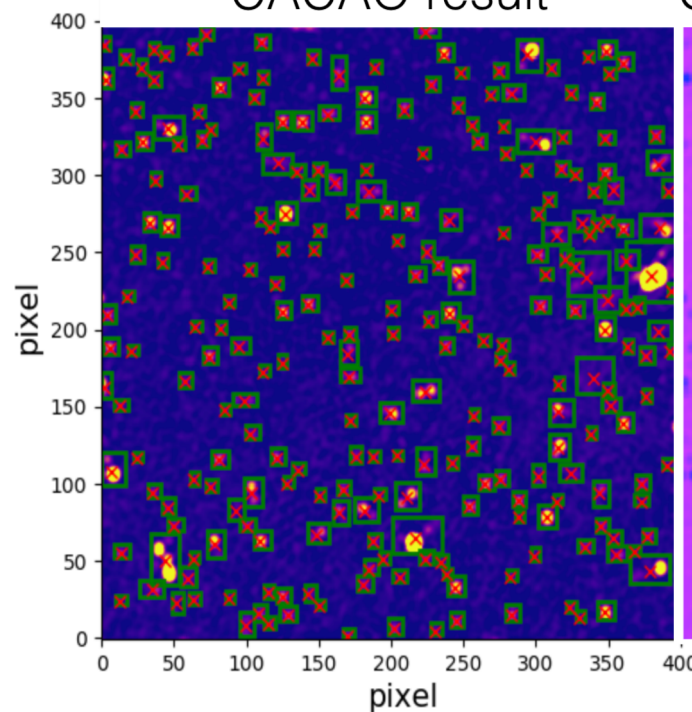
truth table (if available) inspection,
examination of source parameter space,
classification based on source parameters
multi-frequency cross-match using external TOAST code



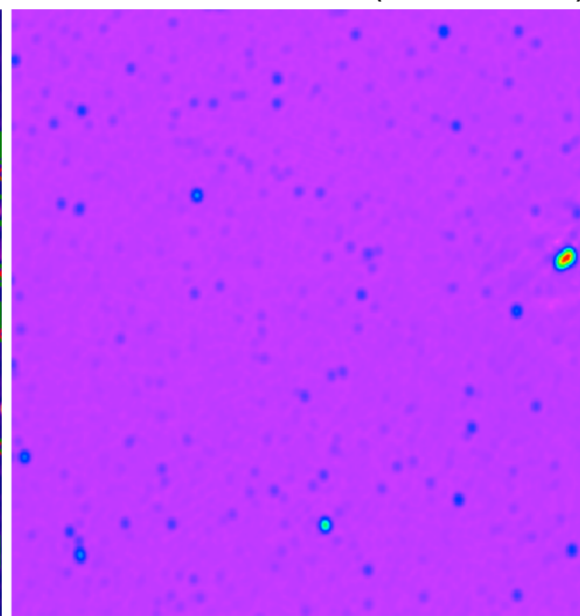
Burkutean et al. in prep

CACAO: The Complete Automated Classification of Astronomical Objects Tool

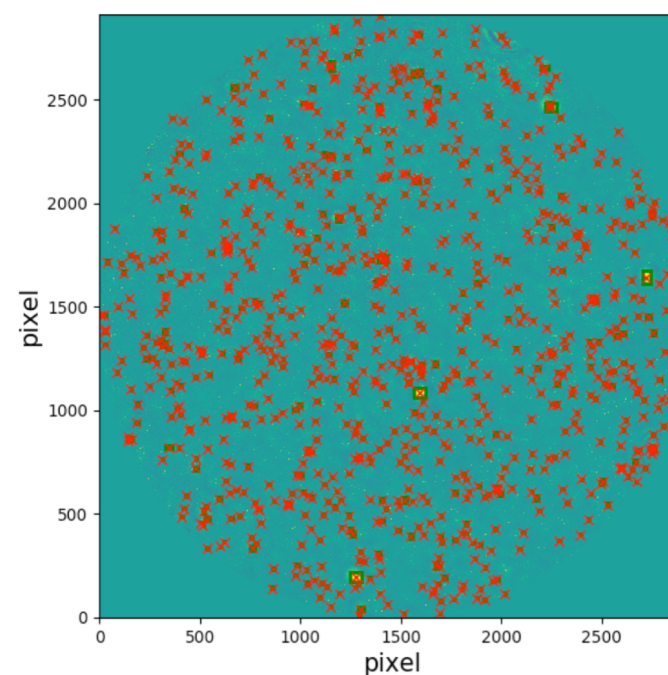
CACAO result



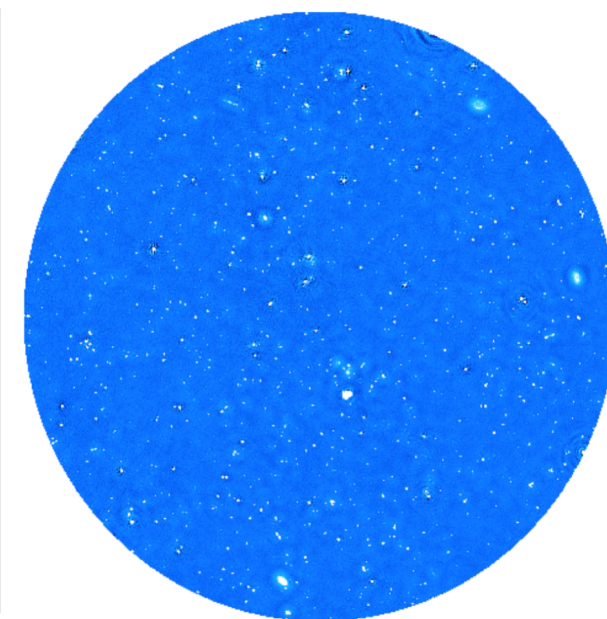
GLEAM cut-out (118MHz)



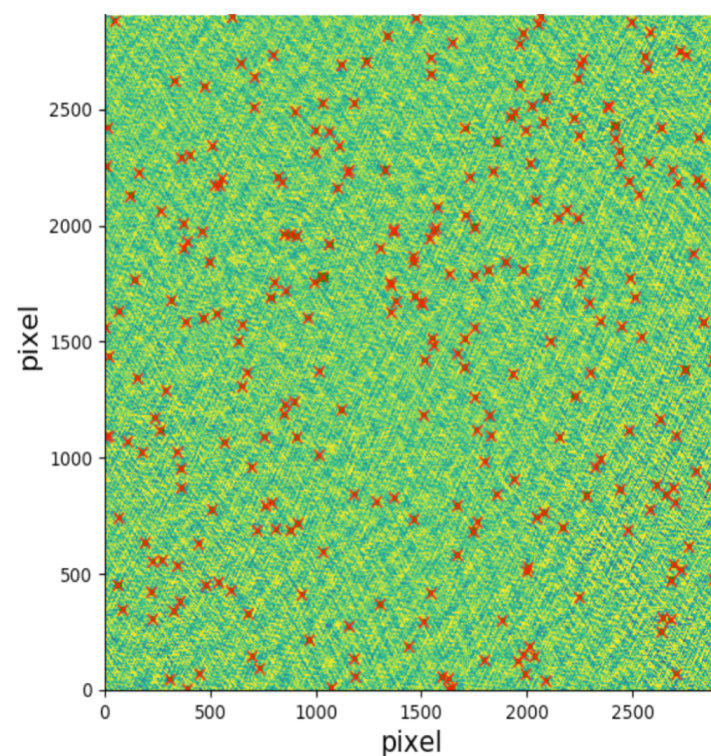
CACAO result



LOFAR, P21-low-mosaic.fits



CACAO result



TGSS cut-out

