Images: Error Recognition and Analysis

Robert Laing
ESO

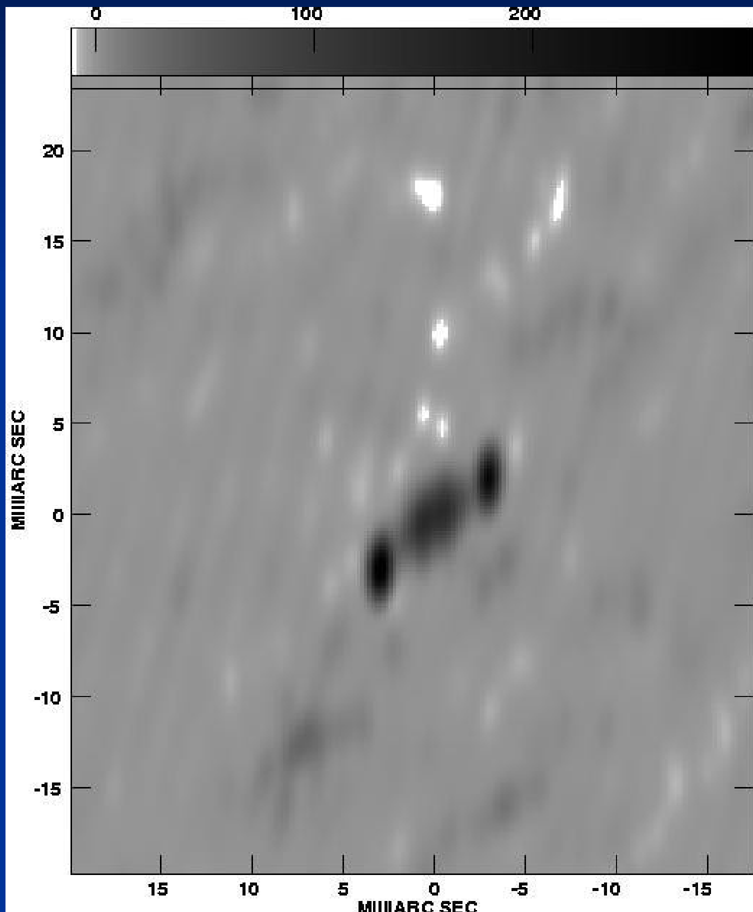
ERIS, ASTRON, Sept 9 - 13 2013

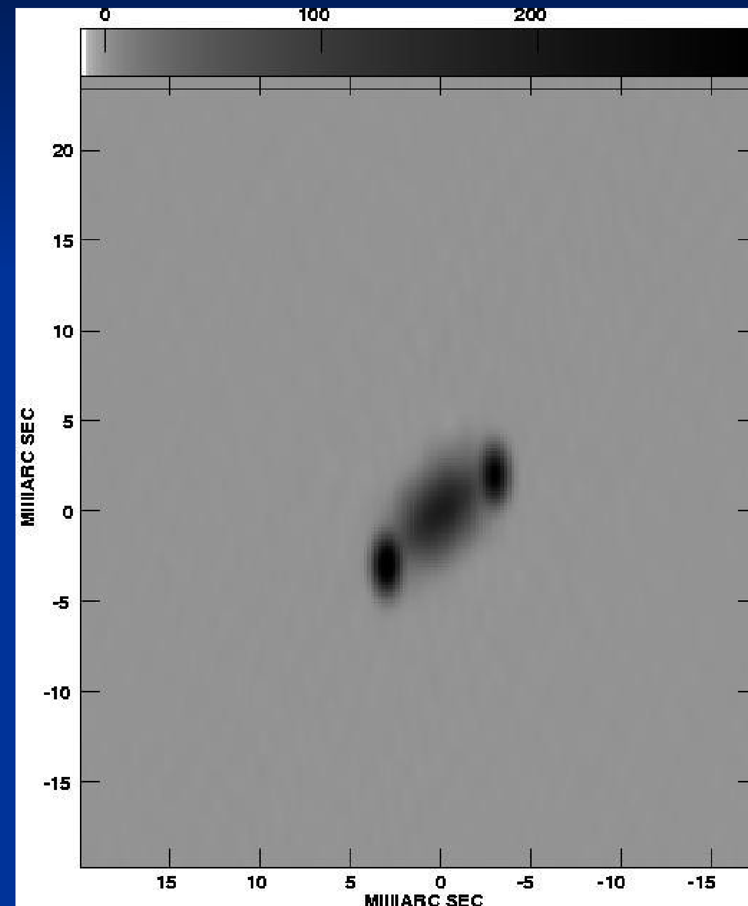
Outline

- Error recognition: how do you recognise and diagnose residual errors by looking at images?
- Image analysis: how do you extract scientifically useful numbers from images

- Unless otherwise specified, this talk is about continuum imaging in full polarization .... but many ideas also apply to spectral-line work.

Have I got a decent image?





No!

Yes!

Robert Laing
ESO

How can I tell (1)?

- Look at the off-source rms:
  - Compare observed off-source rms with that expected theoretically
  - Does the image rms increase near bright sources?
  - Is the noise non-random (e.g. faint waves or ripples)?
  - Compare noise distributions in IQUV

$$S_{rms} = \frac{2kT_{sys}}{A_{eff}\sqrt{N_A(N_A-1)t_{\mathrm{int}}\Delta\nu}}$$

- Look for odd structures
  - Rings, streaks, etc.
  - Coherent features I < -4σ
- Properties of artefacts
  - Additive (constant over the field) or multiplicative (brightest near real structure)?
  - Symmetric or antisymmetric around bright, compact sources?
  - Large or small scale?

How can I tell (2)

- Large-scale negative structures
  - Negative “bowl” around the source
  - Large-scale ripples parallel to source axis
- Small-scale on-source structure
  - Diffuse emission looks spotty
  - High-frequency sine-wave structure across source

Missing short spacings

Deconvolution errors

Possible causes: imaging problems (1)

- Is the image big enough?
  - Aliasing
  - Confusing sources outside the image
- Are the pixels small enough to sample the beam?
  - Are bright point sources accurately located on pixels?
- Wide-field issues
  - Averaging time too long?
  - Spectral channels too wide?
  - w-term?
  - ionosphere?
  - direction dependence of antenna response

Possible causes: imaging problems (2)

- Missing short spacings
- Deconvolution errors, especially with sparse u-v coverage
  - Resolution too high?
  - Poor choice of weighting?
  - Bad choice of CLEAN boxes (too small, too large, ...)
  - Insufficient CLEANing
- Source variability during the observations

Errors in the u-v and image planes

- Errors obey Fourier relations between u-v and image planes
- Narrow features in u-v plane ↔ wide features in image plane and vice versa
- Easier to recognise narrow features
- Orientations are orthogonal
- u-v amplitude errors cause symmetric errors in the image plane
- u-v phase errors cause antisymmetric errors in the image plane

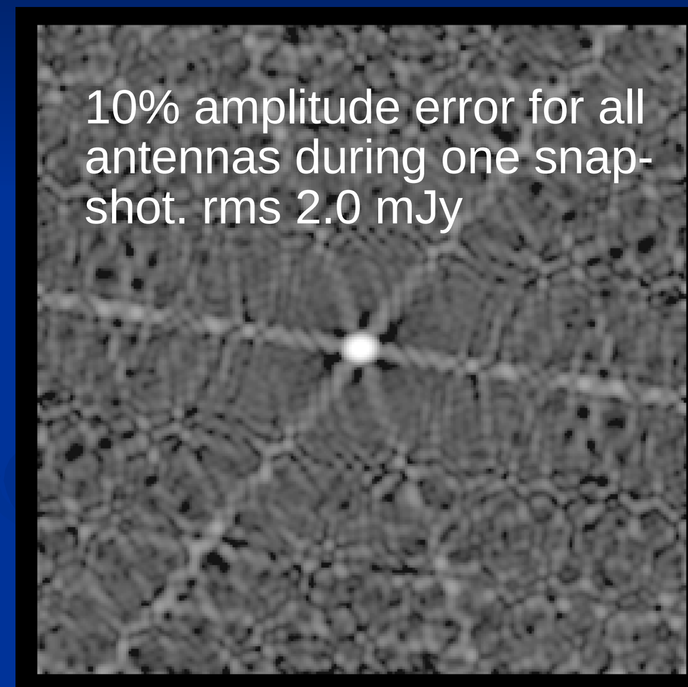
u-v or image plane?

- Find the outliers in the u-v plane first
  - Gross (MJy) points have gross effects on the image
  - A fraction f of bad data points with reasonable amplitudes give fractional error ~f in the image
- Low-level, but persistent errors are often easier to see in the image plane
- Rule of thumb: 10 deg phase error ≡ 20% amplitude error

Amplitude errors: all antennas

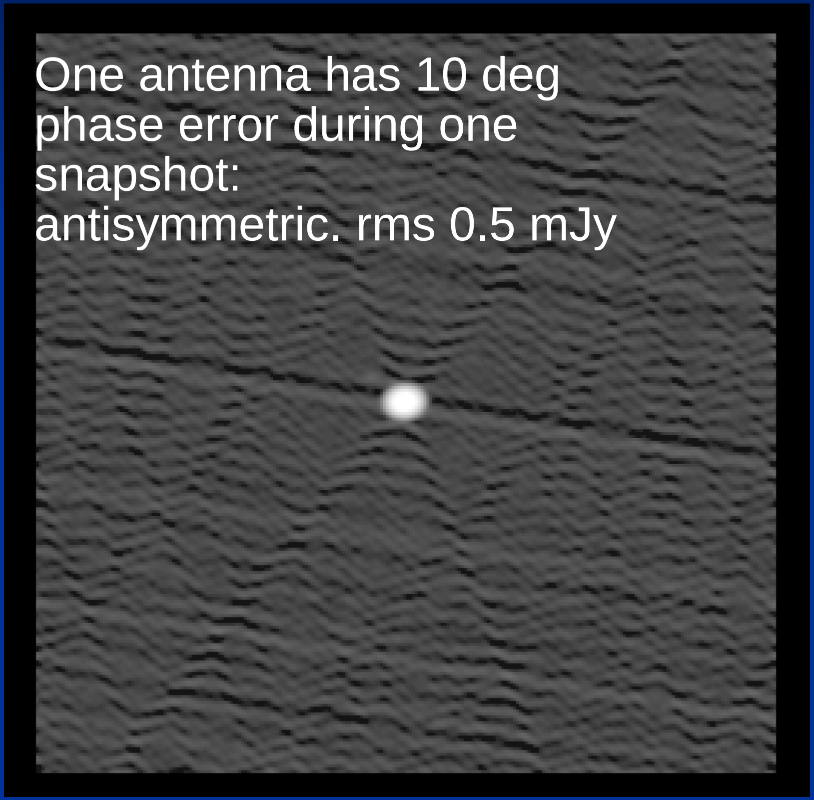
VLA: point source, 13 x 5 min snapshots over 10 hours

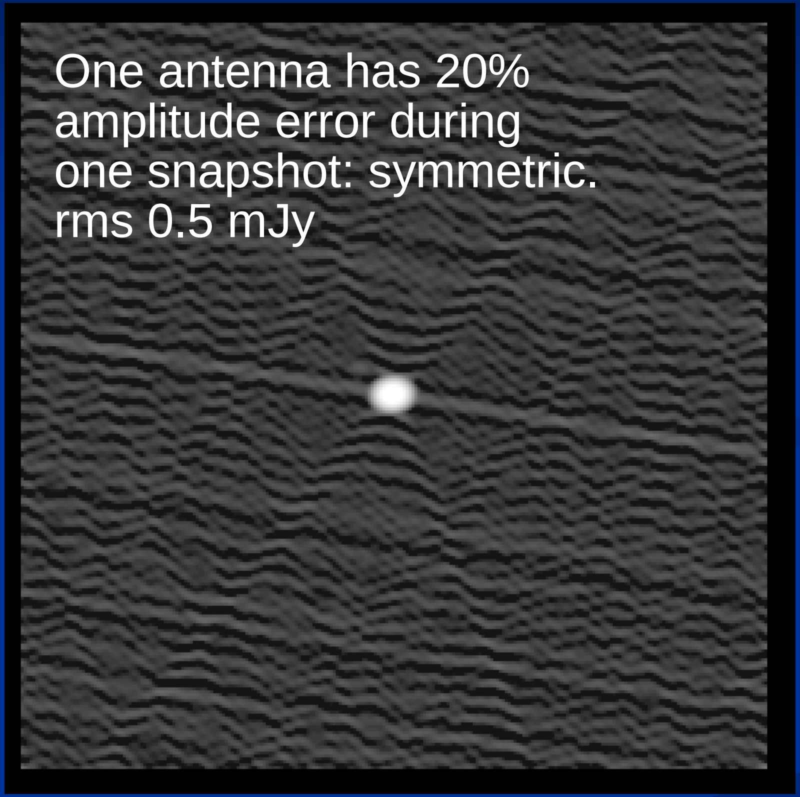
N.B. error pattern looks like the dirty beam



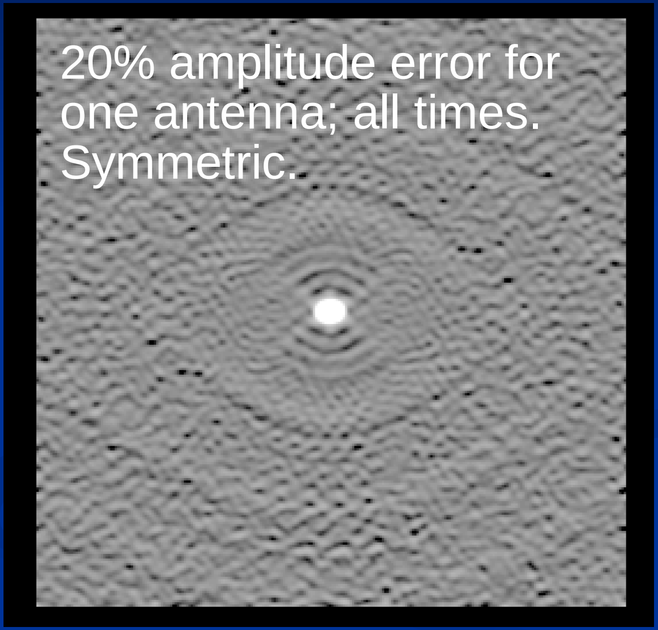
Multiplicative

One antenna in error at one time





One antenna in error: all times

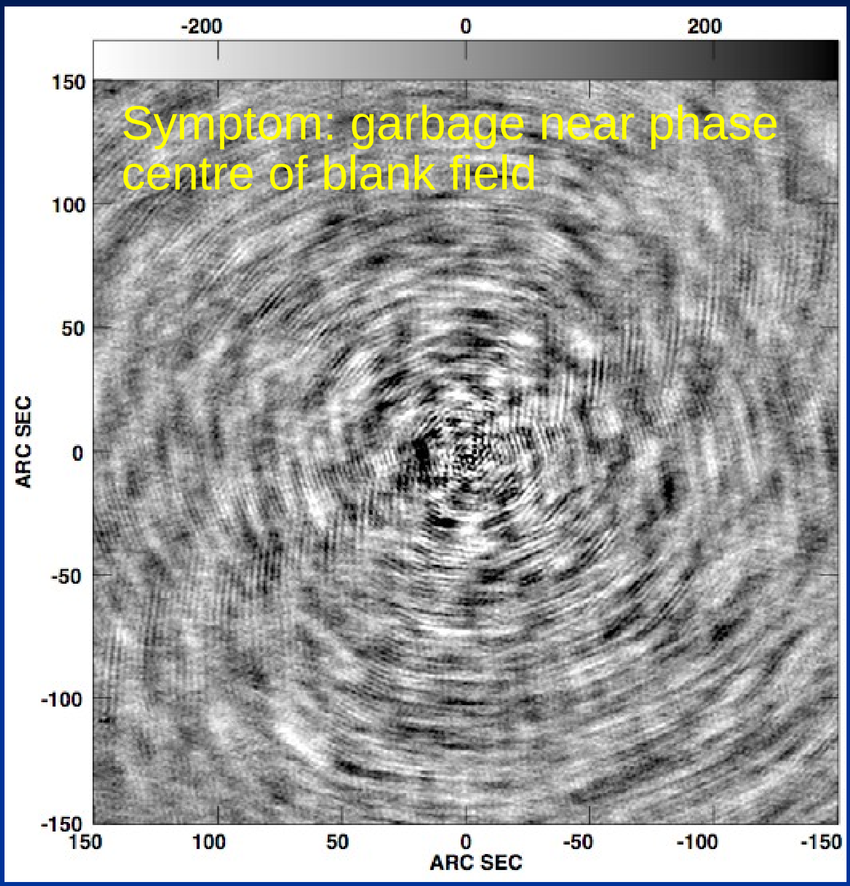


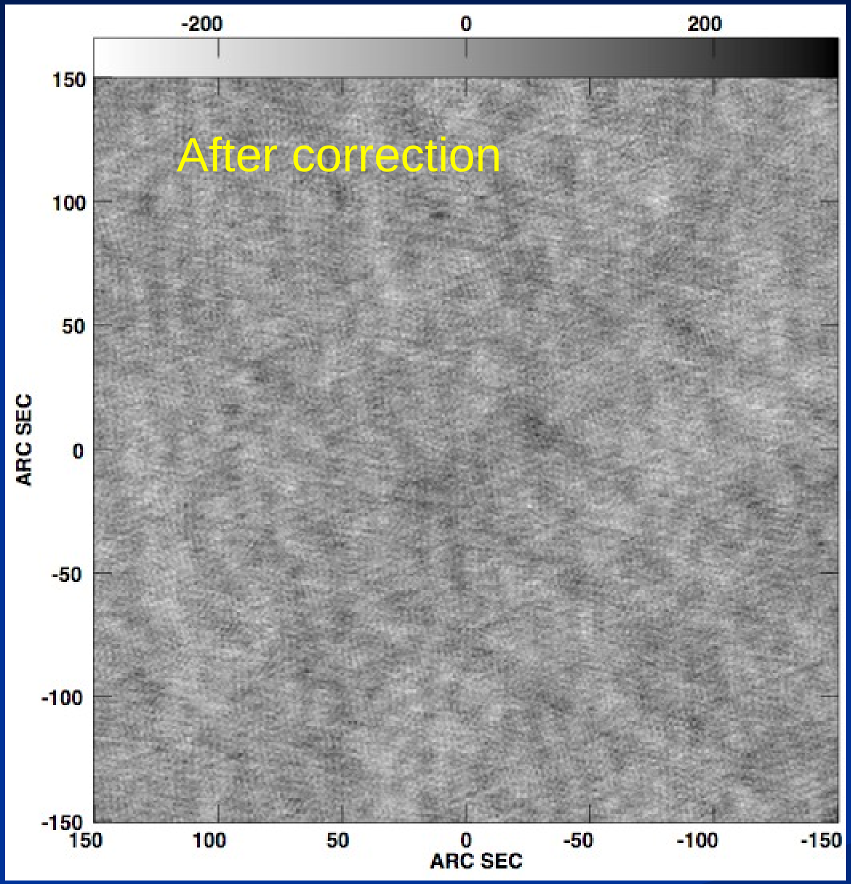
Multiplicative

Multiplicative

Can diagnose by dropping one antenna in turn and re-imaging

Correlator offsets

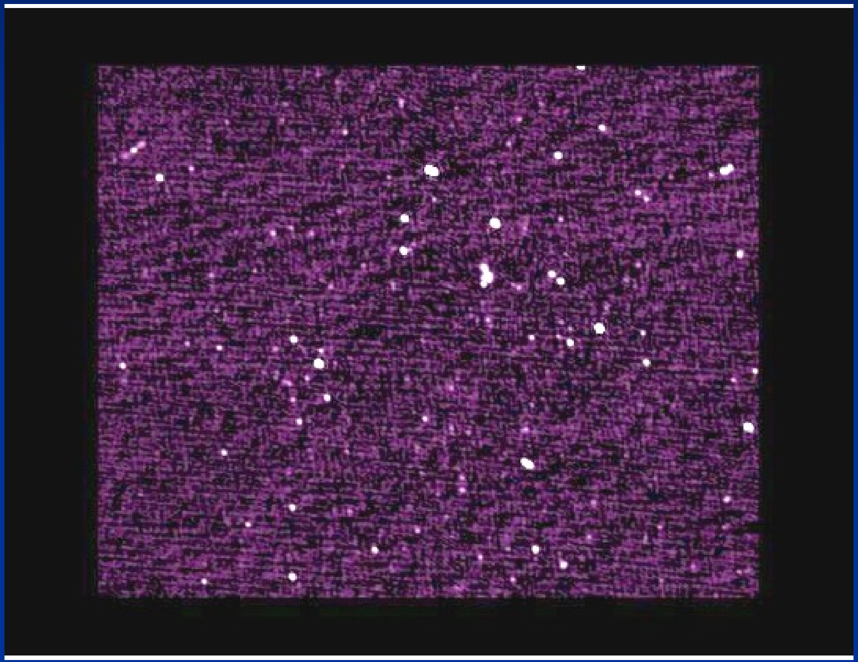




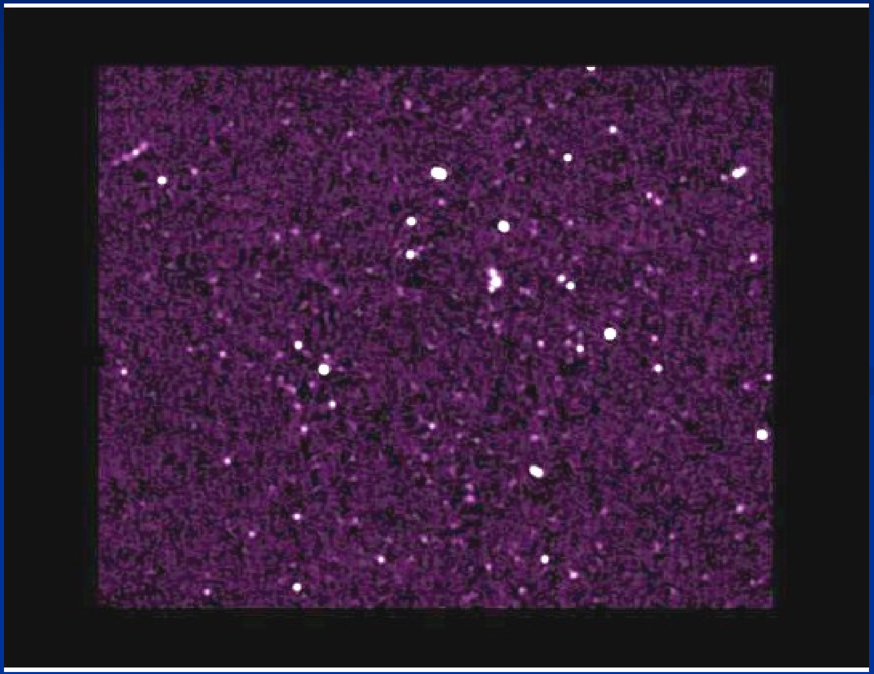
Additive

Non-closing additive errors on individual baselines

Finding subtly bad data



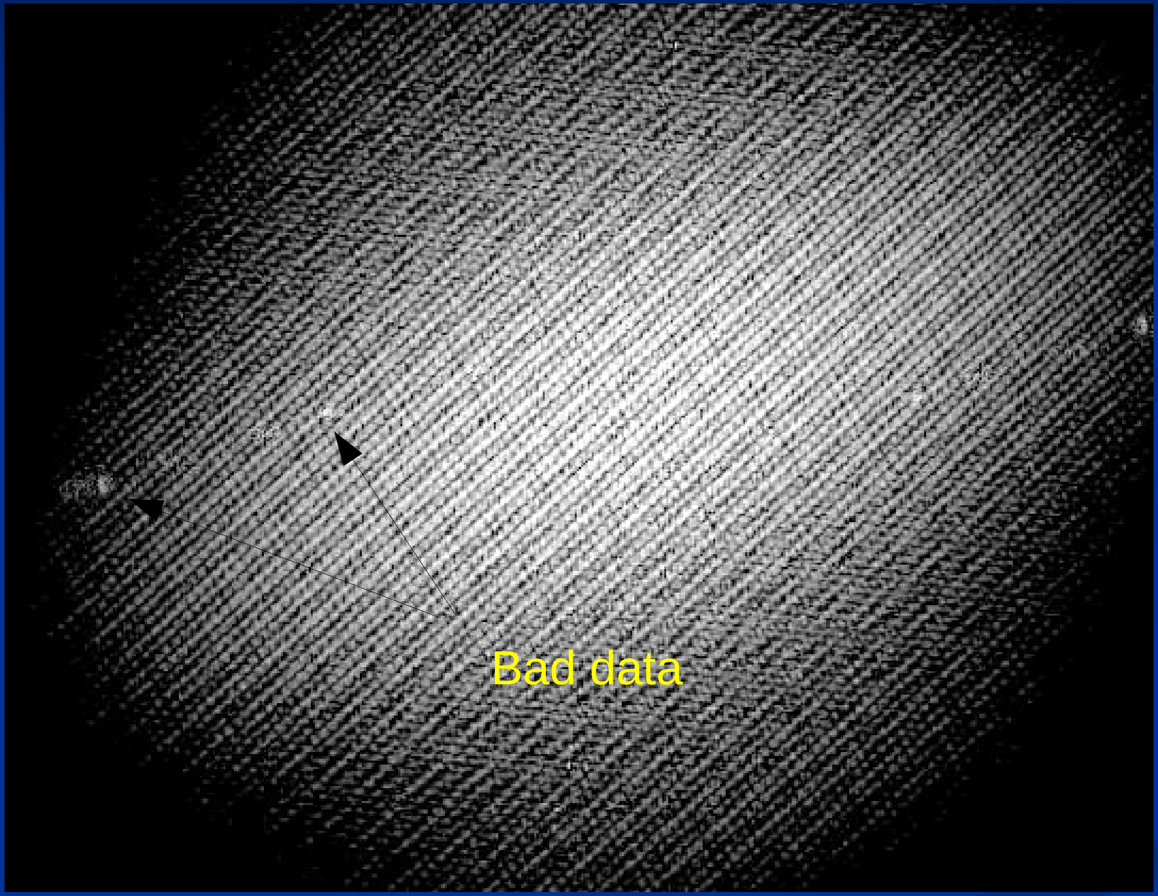
Before editing



After editing

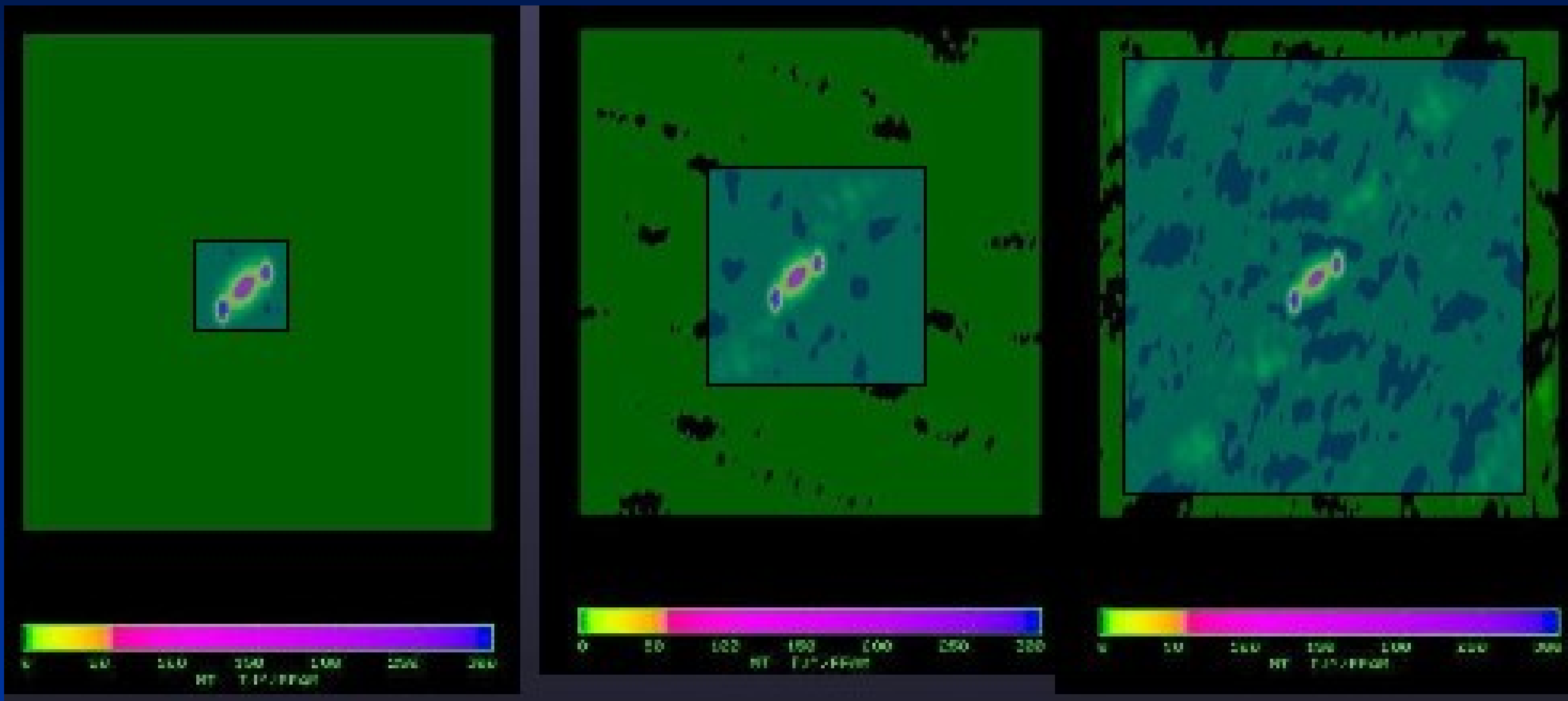
CDFS rms 0.02 mJy

Fourier transform of CLEAN image



Problem is incorrect weighting of a small number of visibilities

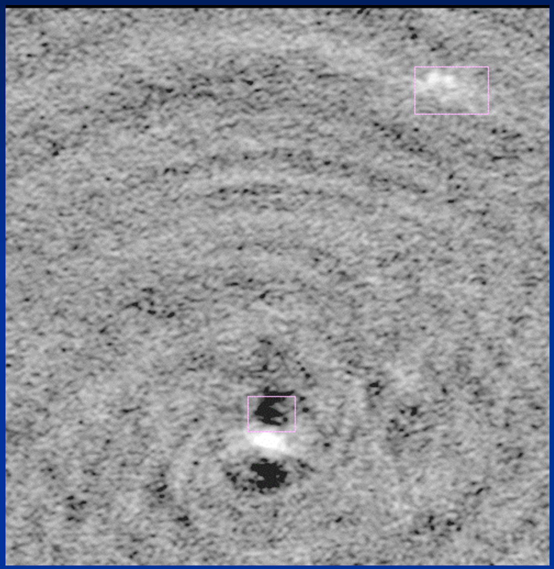
Window cleaning

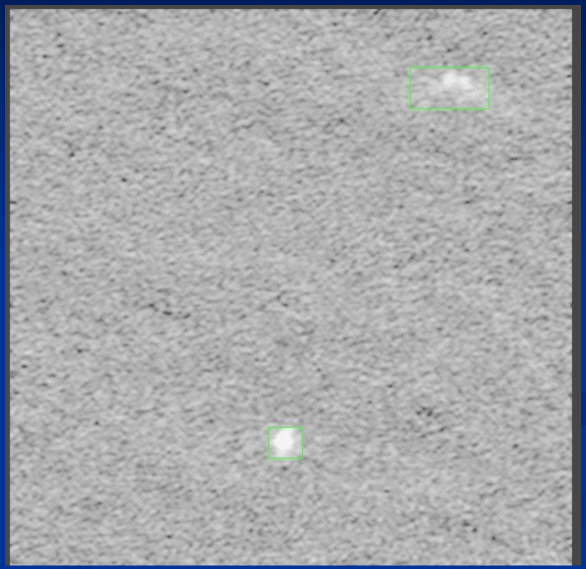


Making size of CLEAN box (AIPS) / mask (CASA) too big leads to larger residuals, typically with the same shape as the dirty beam – but be careful to include all of the real emission

CLEAN functions best if the area over which it is allowed to find components is restricted - fewer unknowns; same number of constraints ('compact support')

Under-cleaning

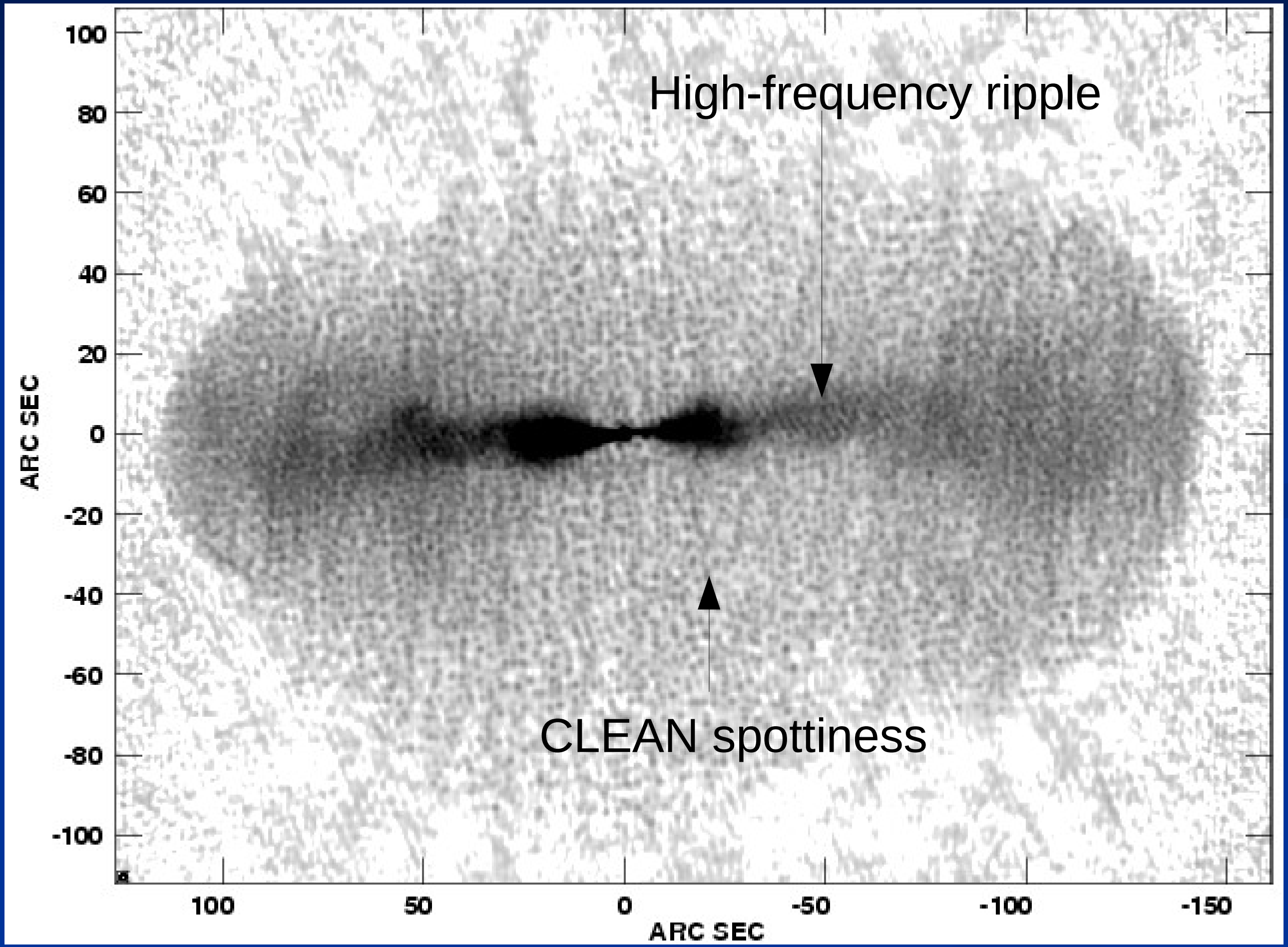




Insufficient cleaning   Adequate cleaning

Over-cleaning can also produce artefacts, especially for poorly-sampled data

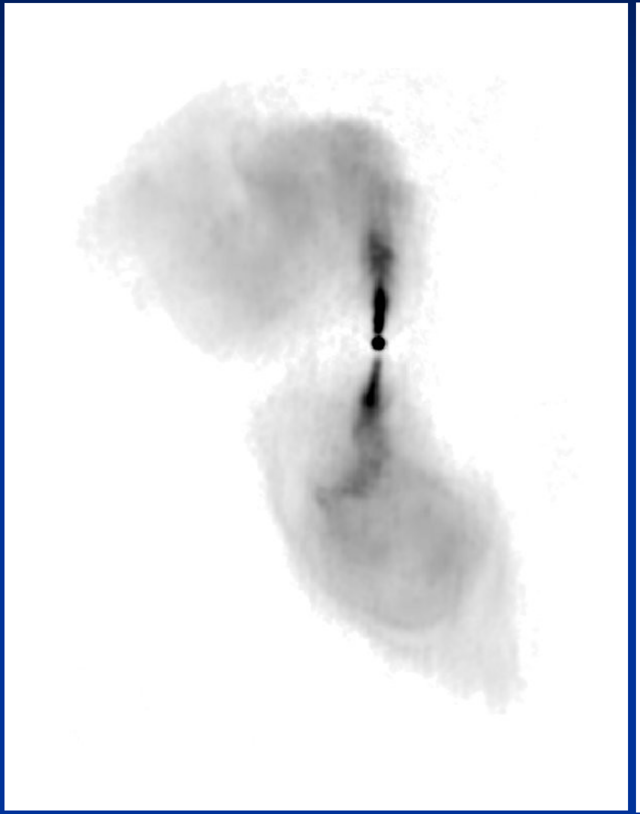
Deconvolution problems



VLA A+B+C configurations. Short spacings OK, but with poor A-configuration coverage

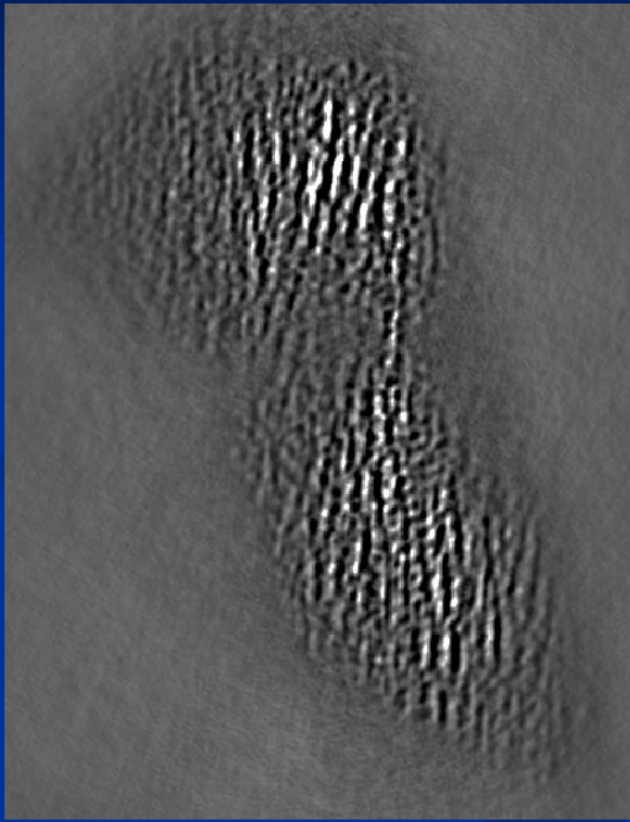
Conventional CLEAN fails: try multi-resolution CLEAN or MEM or reduce the resolution

Multi-scale CLEAN helps



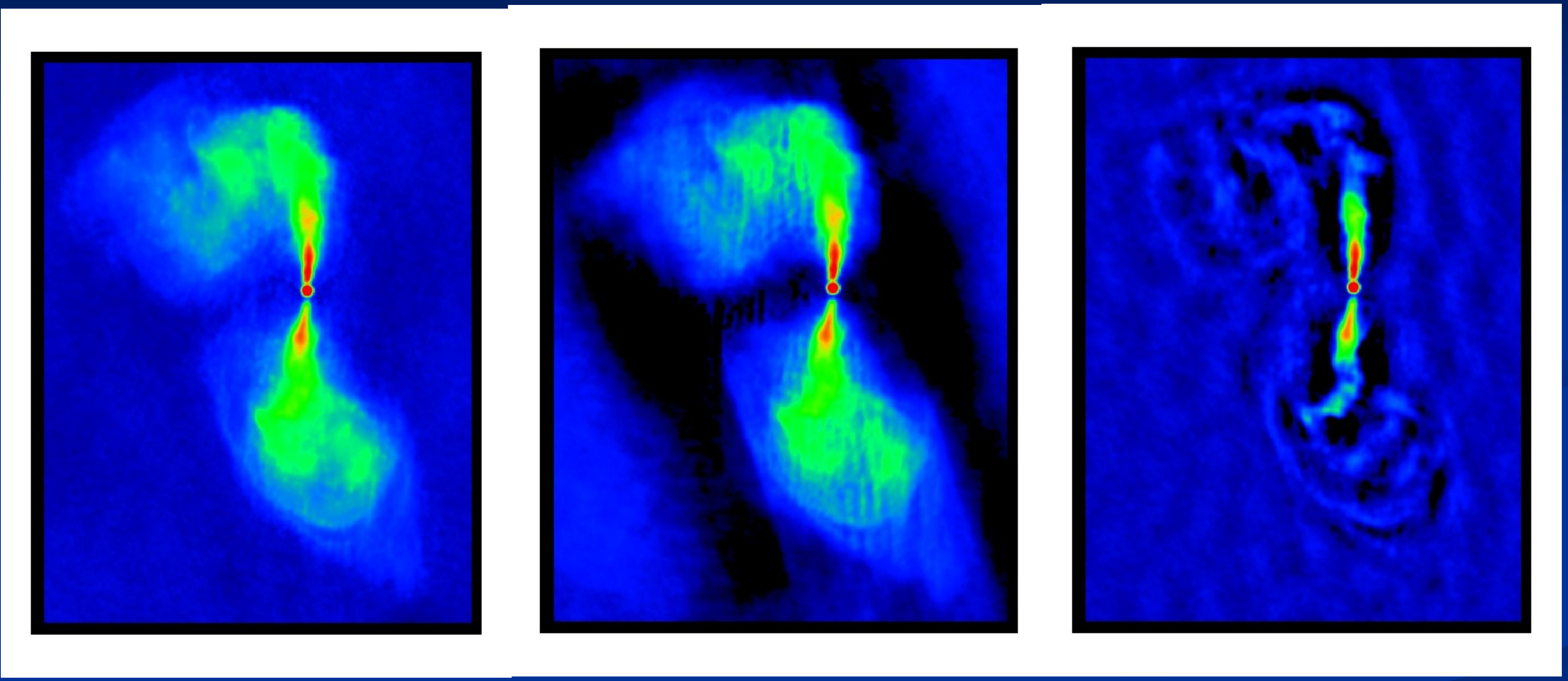
3-scale CLEAN

1-scale CLEAN



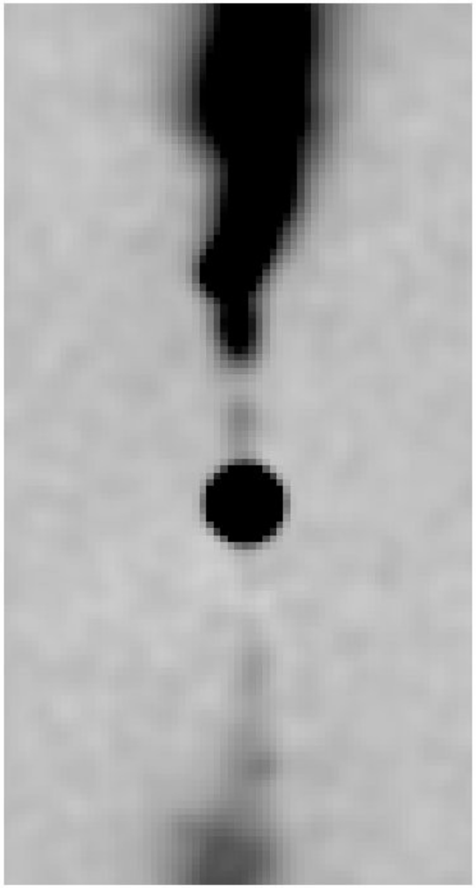
1-scale - 3-scale

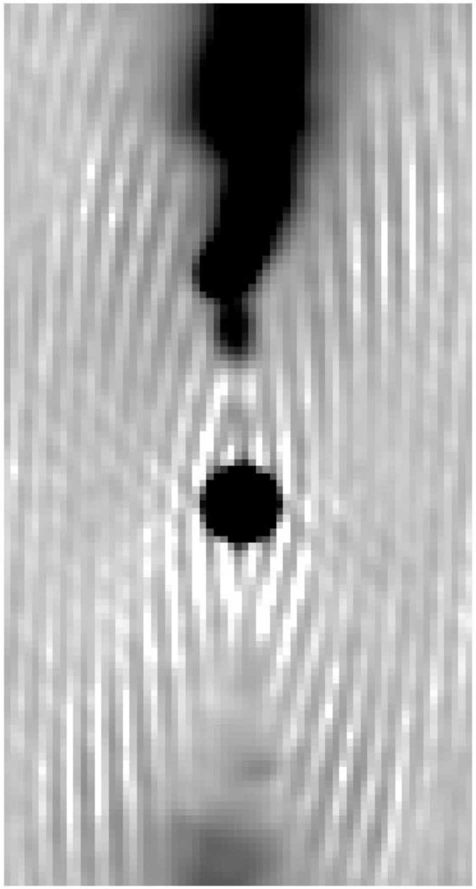
Missing short spacings



uv range < 225 kλ   uv range 2 – 225 kλ   uv range 10 – 225 kλ

Point source not on a pixel

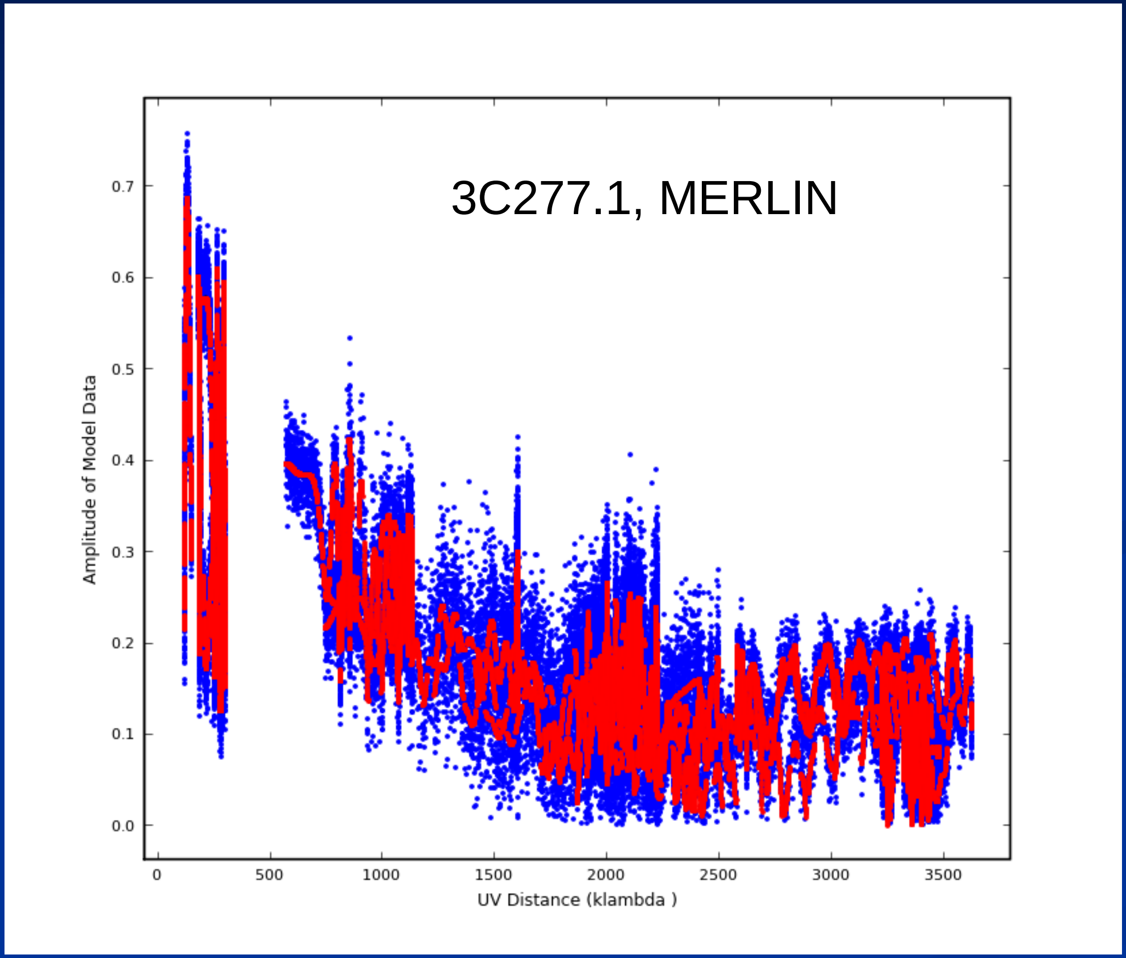




Point source centred

Point source half way between pixels

Does the model fit the data (1)?

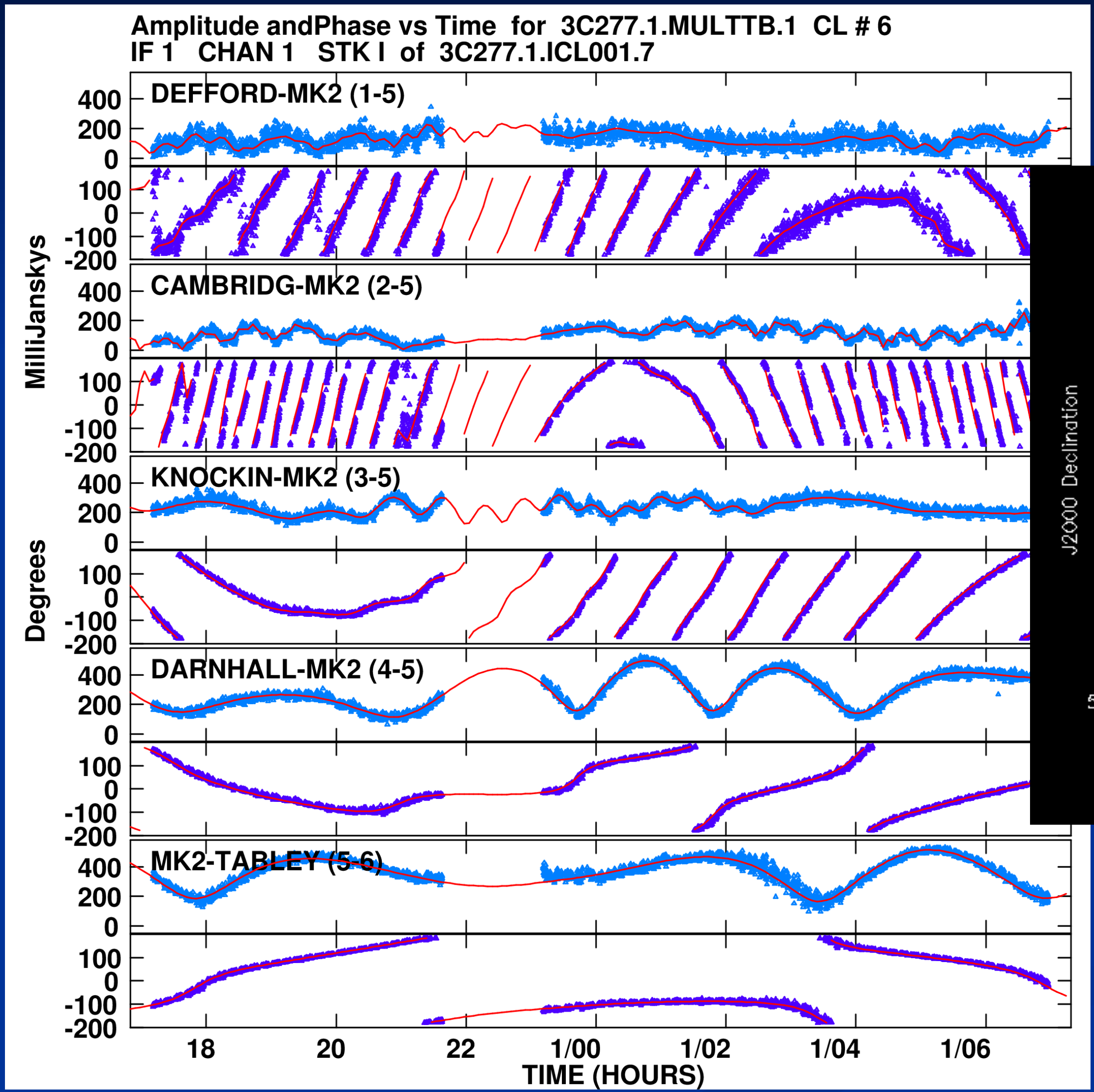


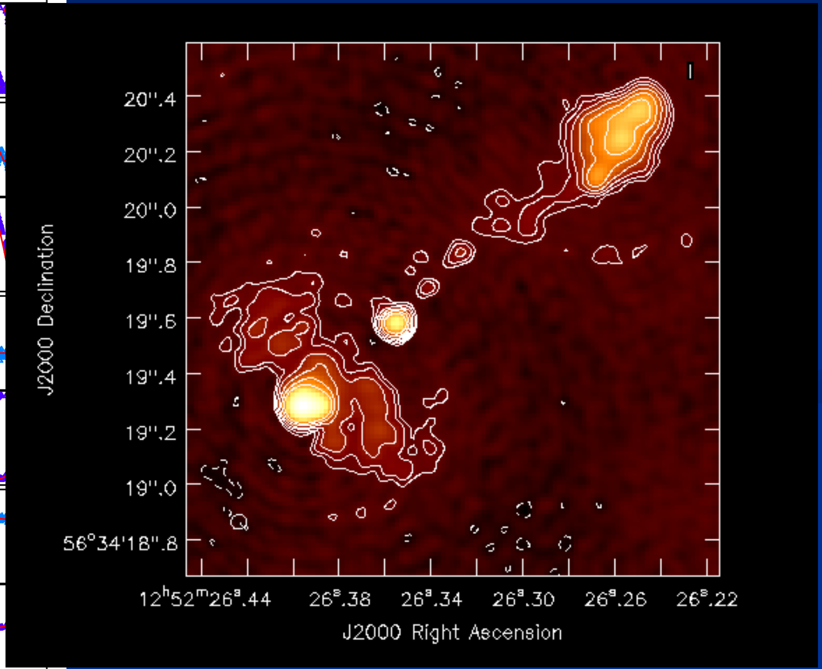
Plot amplitude against uv distance

Data

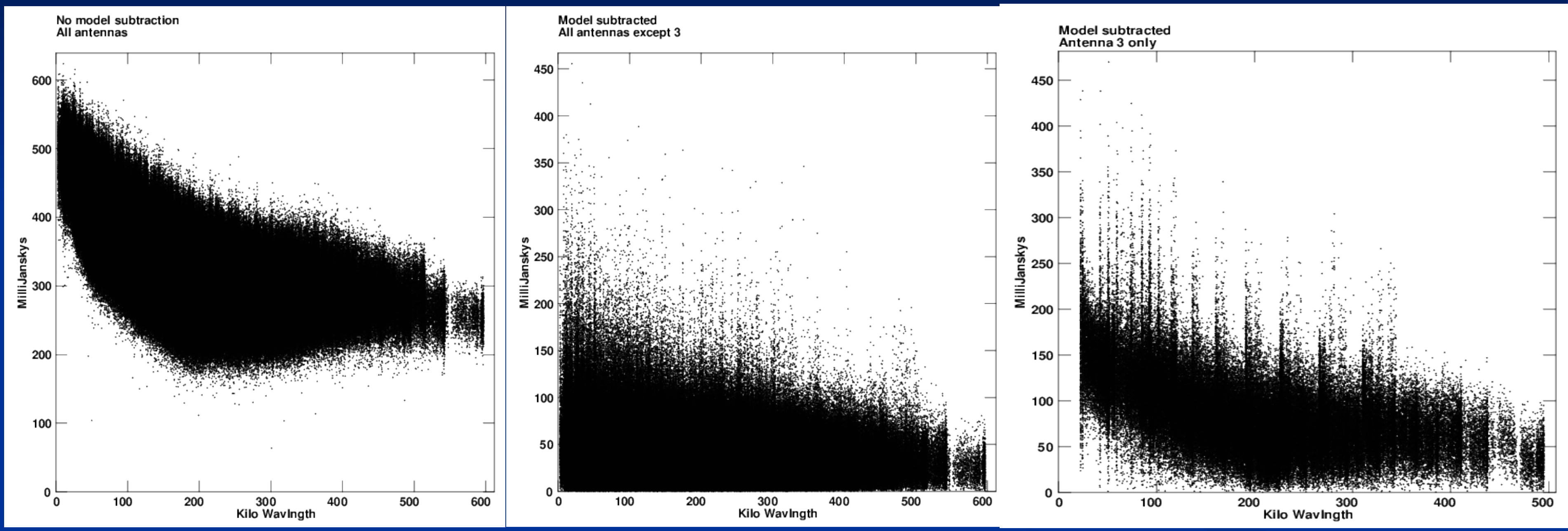
Model

Does the model fit the data (2)





Does the model fit the data (3)?



Error present; all antennas

Model subtracted; all antennas except 3

Model subtracted; antenna 3 only

Persistent phase error in antenna 3 only

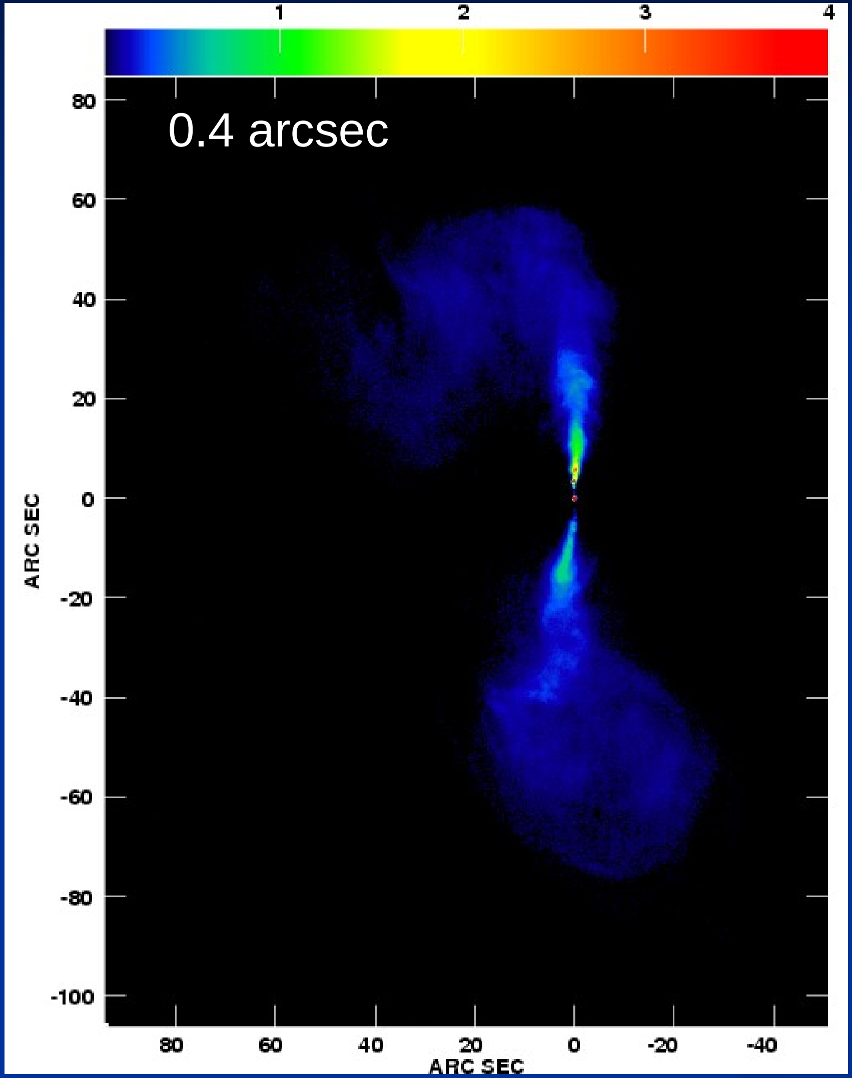
Summary of error recognition

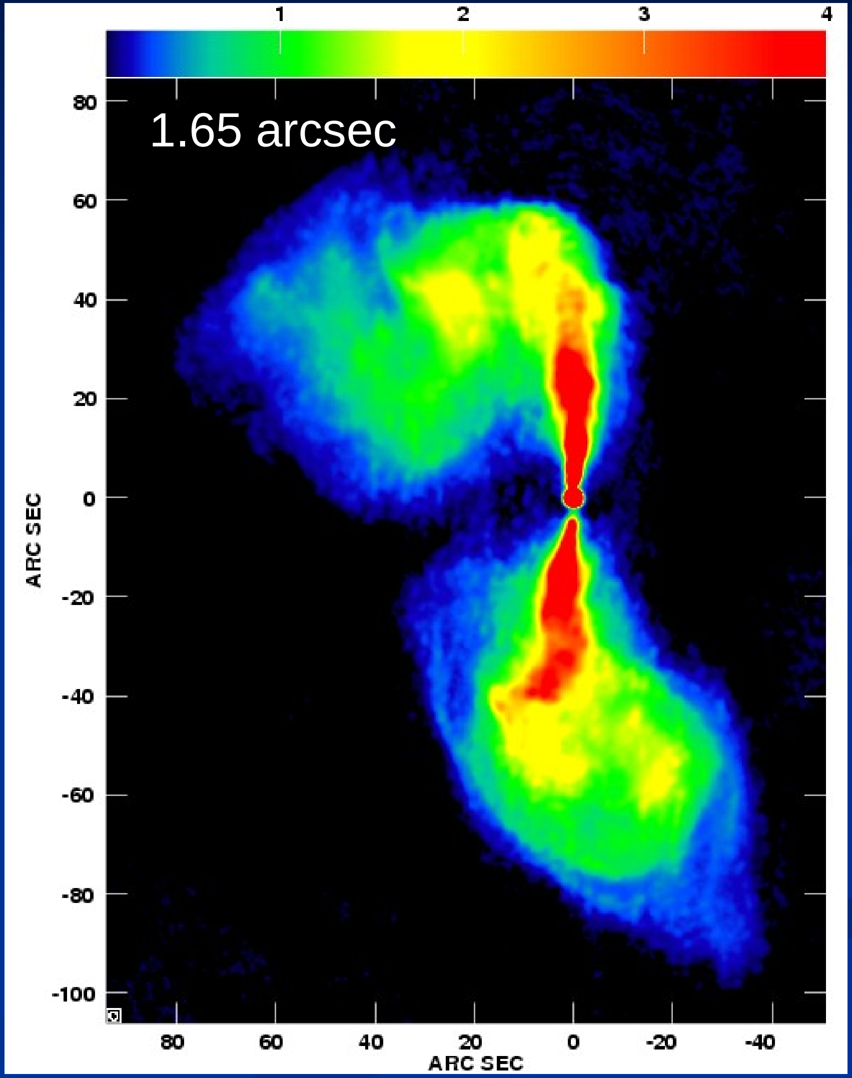
- u-v plane
  - Look for outliers
  - Check gains and phases
  - Look for residuals (data – model)
- Image plane
  - Do the defects look like the dirty beam?
  - Additive or multiplicative?
  - Symmetry properties?
  - Relate to possible data errors
  - Deconvolution problems
- If in doubt, simulate with realistic error model
  - CASA simulation tasks
  - ALMA Observation Support Tool

Image analysis

- Given: a well-calibrated dataset producing a high-quality image (or, in general, image cube)
- How can we extract scientifically useful numbers?
- This is a very open-ended problem, depending on:
  - image complexity
  - scientific goals
- Selected topics (excluding spectral line):
  - Picking the correct resolution
  - Parameter estimation
  - Comparing images: spectra, polarization etc.; registration
  - Getting images into your own code

Match the resolution to the problem

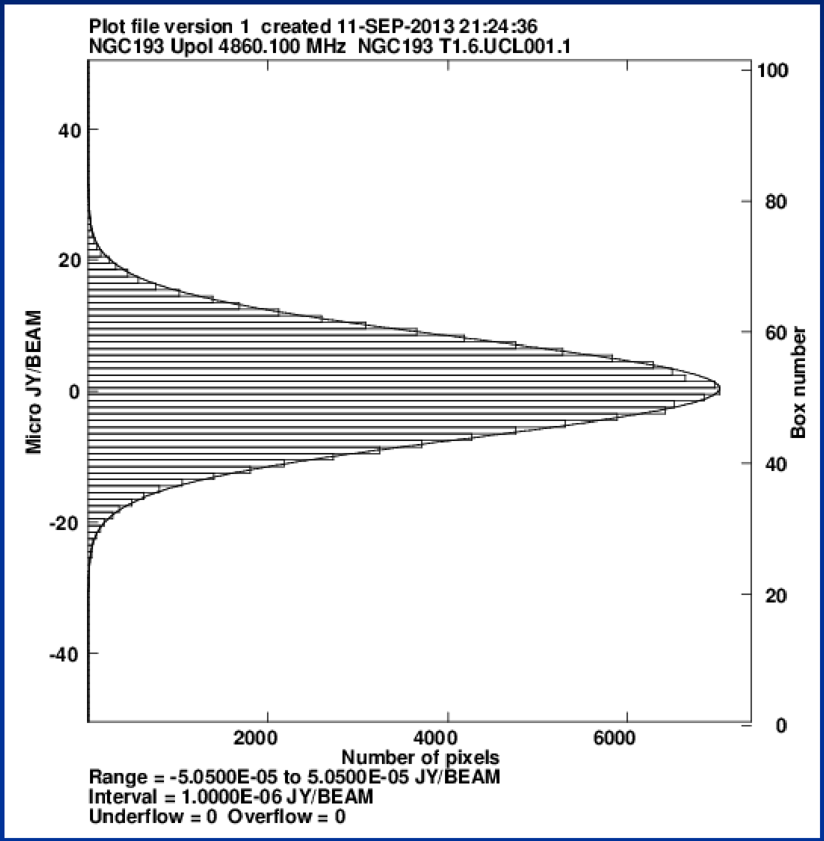


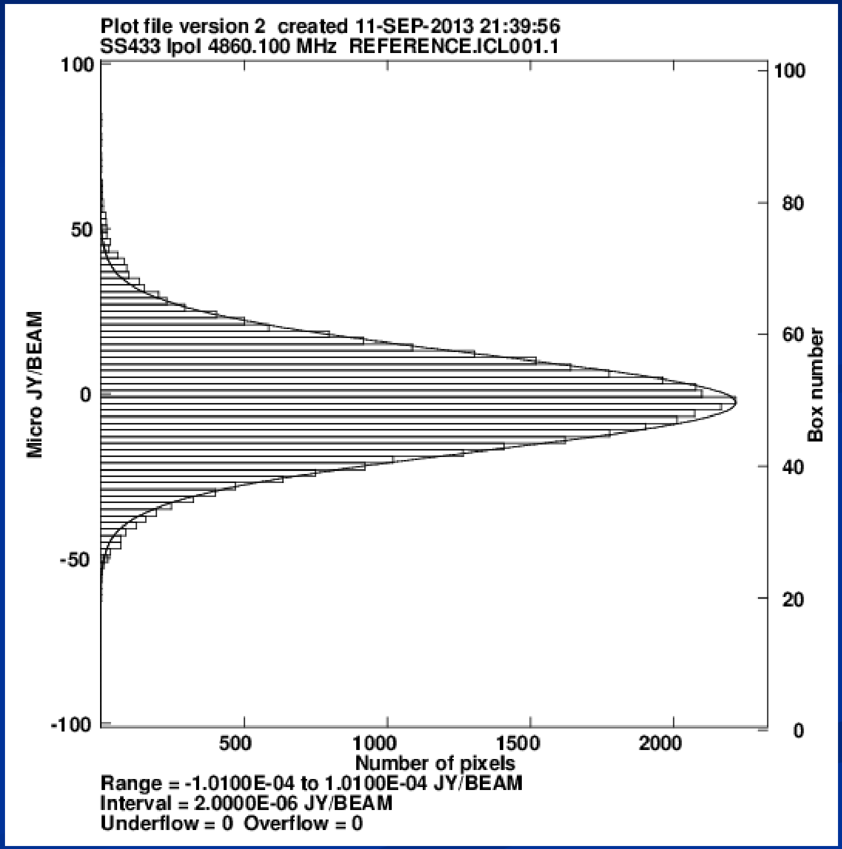


Same dataset, imaged with different Gaussian tapers

Noise Distributions

- Measure rms (and mean) over off-source regions





Good case: rms = 7.5μJy; Gaussian noise with zero mean

Excess noise above Gaussian tail

Estimating integrated flux density of an extended source

- Use a **low-resolution** image, cleaned deeply
  – The beam areas of the restored CLEAN components and residuals are not the same in general.
- Sum the flux density over some area (rectangular, polygonal, ...) – AIPS IMSTAT, TVSTAT; CASA imstat, viewer.
- Remember that the total flux density is ΣI/B, where B is the integral over the beam. For a Gaussian,

$$B = \pi(\mathrm{FWHM/pixel})^2/4 \ln 2.$$

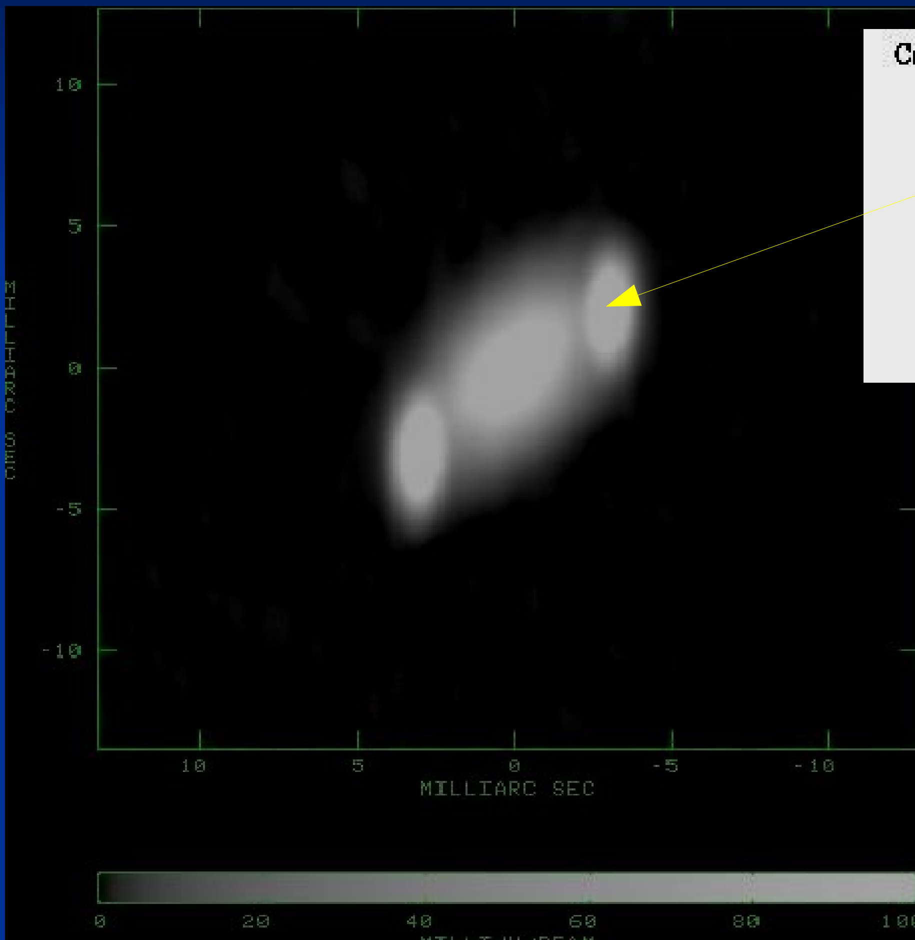
The reduction packages will calculate this for you.

- The reason is that the images are normalised so that a point source of flux density 1 Jy gives a **peak** response of 1 Jy/beam on the image.

Component fitting

- Image plane
  - Assume source components are ~Gaussian
  - Deep cleaning restores images with Gaussian beam
  - Size estimation quite straightforward
- u-v plane
  - More accurate for small numbers of ~point-like sources
  - Can fit to models that are slightly more complex than point-like
  - Accounts for imperfect sampling; noise distribution may be better understood, but ...
  - ... no good for complex brightness distributions
- Error estimates
  - ad hoc
  - From fitting routines
  - By simulation

Gaussian fitting example




```
Component   1-Gaussian
  Peak intensity     = 0.300 +/- 0.005 JY/BEAM
  Integral intensity= 0.302 +/- 0.008 JANSKYS
  X-position         =     270.991 +/-    0.001 pixels
  Y-position         =     267.018 +/-    0.001 pixels
    Major ax               0.53    +/-    0.01  pixels
    Minor ax               0.00    +/-    0.05  pixels
    Pos ang               21.6     +/-    1.1   deg
```

AIPS JMFIT
CASA imfit

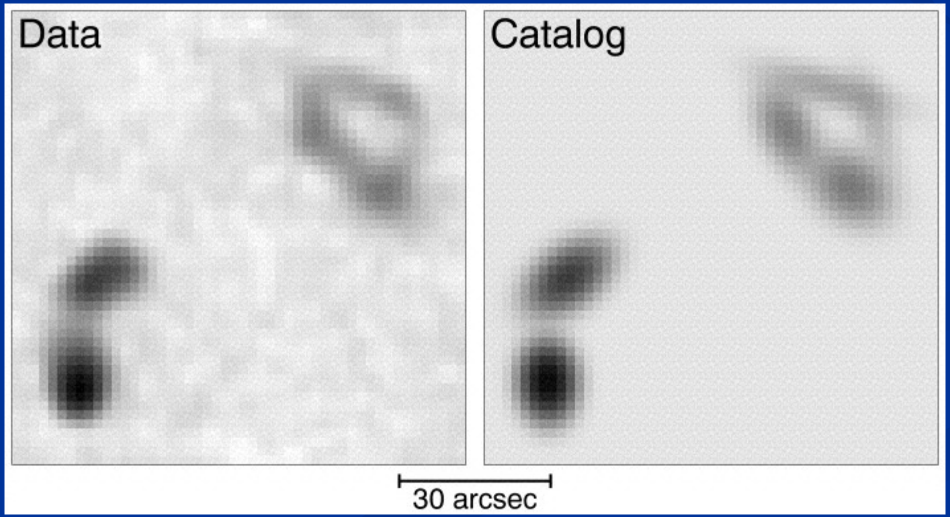
Errors

A priori error estimates

- Definitions
  - P = peak component flux density
  - σ = image rms noise
  - θ_B = CLEAN beam size
  - θ_{obs} = component size
  - S = P/σ = signal/noise
- rms errors
  - Error on peak flux density = σ
  - Position error = $\theta_B/2S$
  - True component size $\theta = (\theta_{obs}^2 - \theta_B^2)^{1/2}$
  - Minimum measurable component size = $\theta_B/S^{1/2}$
    - S/N >100 is needed to determine a size $<\theta_B/10$.

Automated image-fitting

- Automated routines (e.g. SAD) can be used to locate and fit sources (essential for surveys). Available in standard radio-astronomy packages.
- Also adapt routines used in optical astronomy (e.g. Sextractor)
  - beware incorrect noise model
- Often worthwhile to make Monte Carlo simulations to assess realistic errors in position and (especially) flux density (e.g. add model point sources).

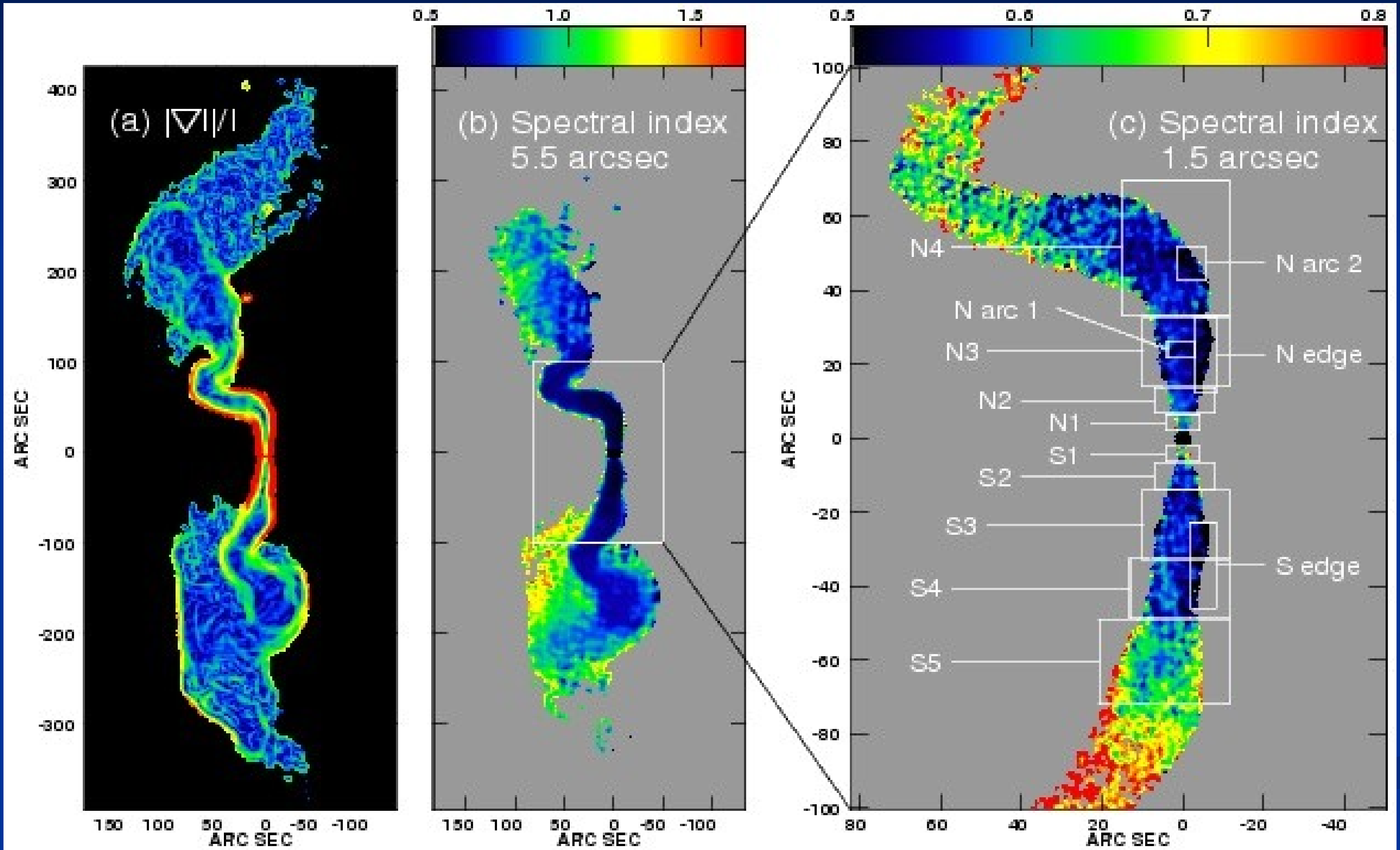


Output of automated Gaussian component fitting as used in the FIRST survey (White et al. 1997)

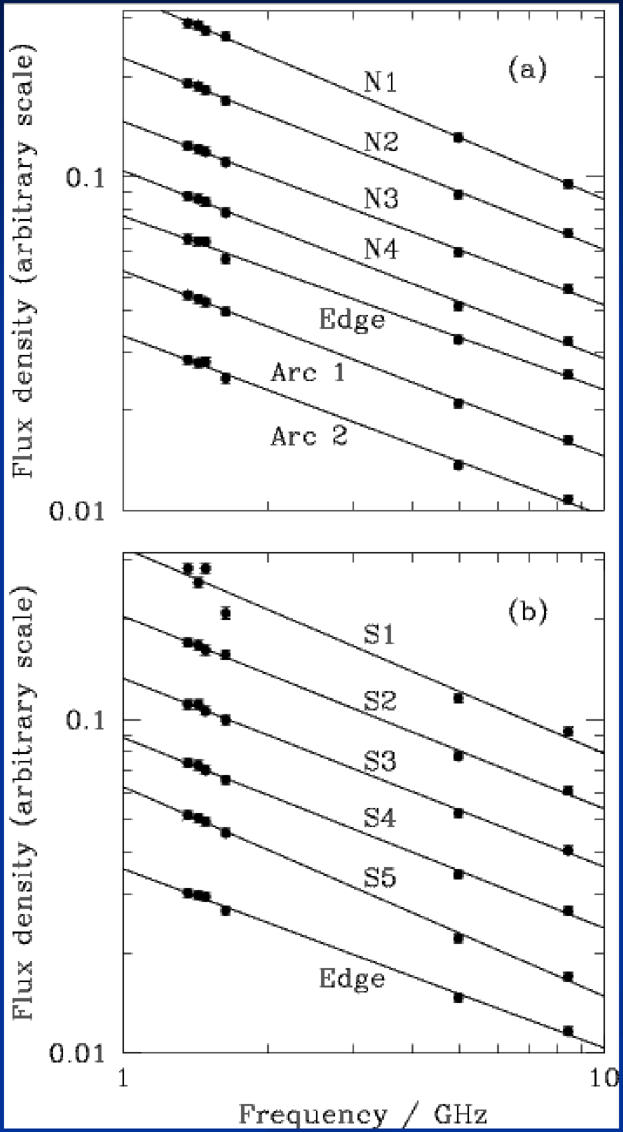
Image arithmetic

- Standard packages allow mathematical operations on one, two or occasionally more images (AIPS MATH, COMB, RM; CASA immath):
  - Sum, product, quotient, ...
  - Polarized intensity and position angle from Q and U
  - Spectral index α ($S \propto \nu^{\alpha}$)
  - Faraday rotation measure
  - Optical depth
  - ....
- Can also propagate noise and blank on input values or s/n
- Other image manipulations (spatial filtering, etc.) are also possible
- Current packages are poor at fitting functions of frequency to images at more than 2 frequencies: usually have to do this yourself.

Spectral index and Gradient filter



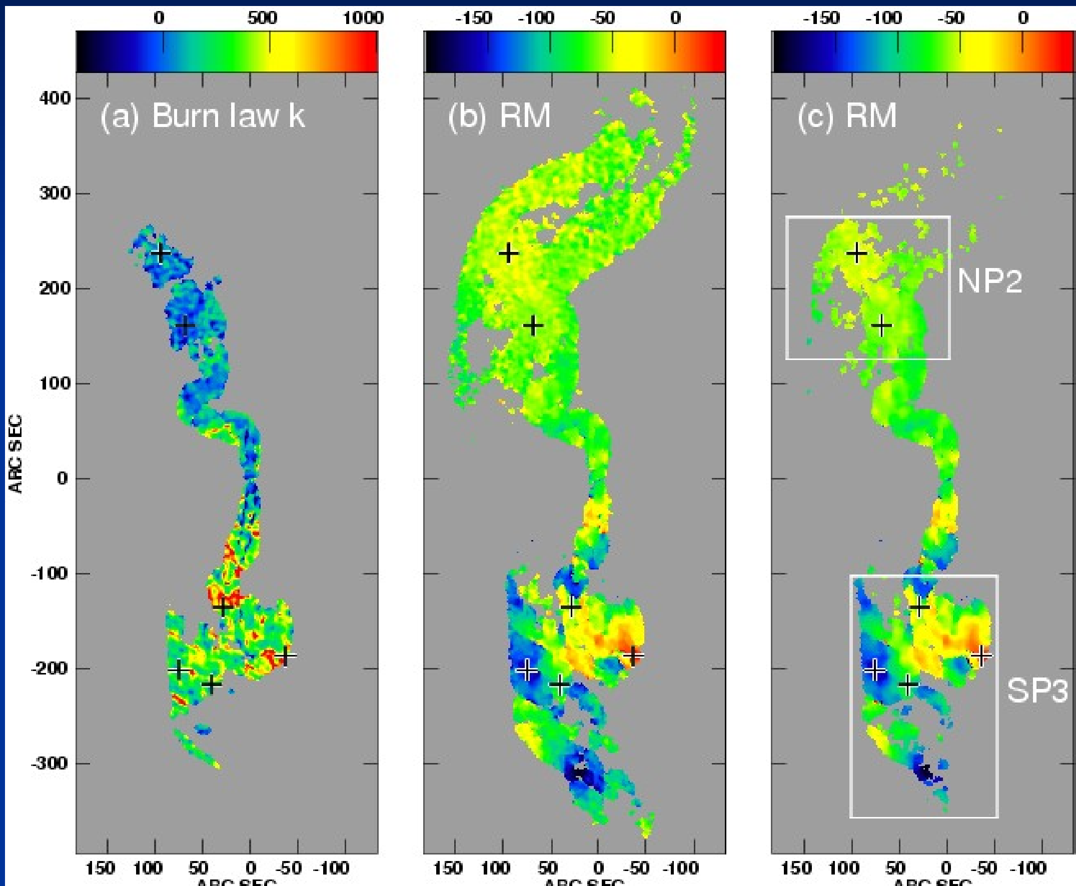
Spectra

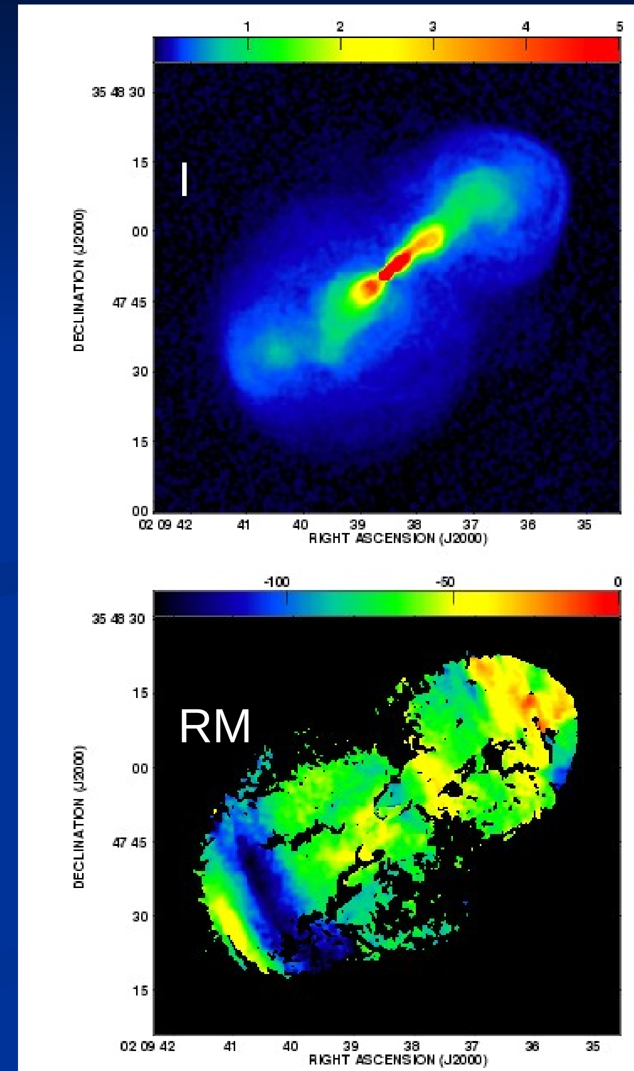


Spectra derived by integrating the flux densities over the boxes shown on the previous slide.

Note the slight flux scale error at 5 and 8.4 GHz.

Rotation measure and depolarization





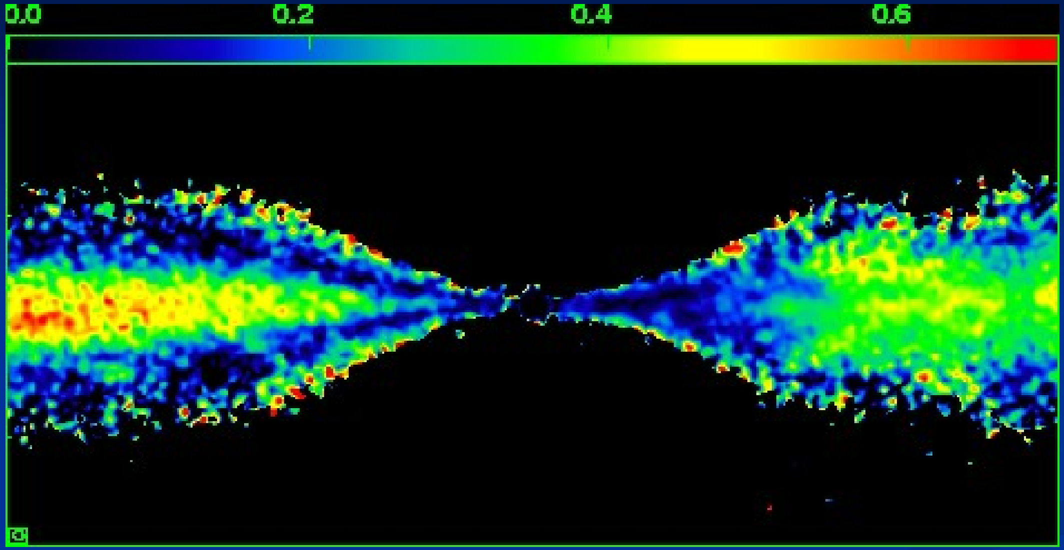
5 – frequency rotation measure images

$\chi = \chi + RM \lambda^2$

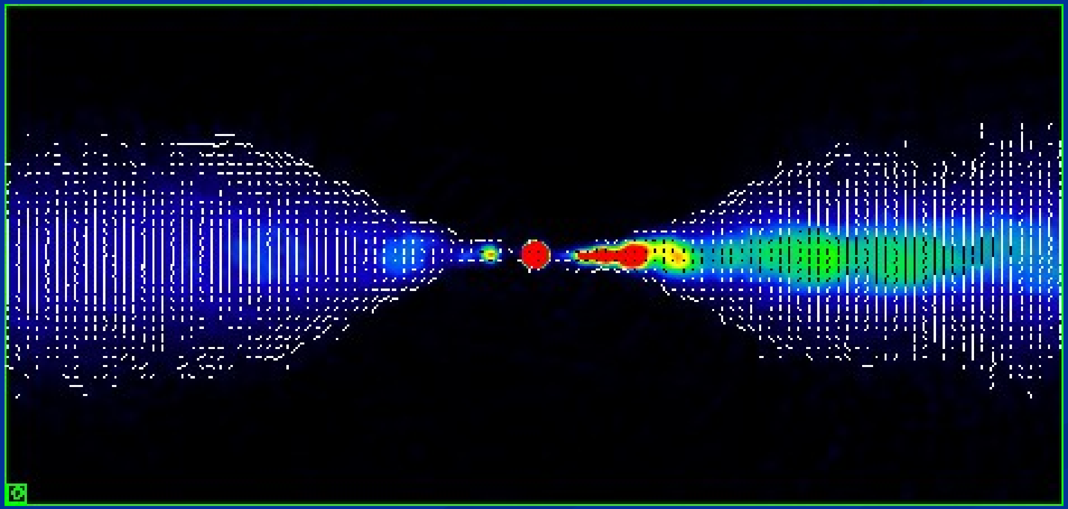
(Laing et al. 2008

Guidetti et al. 2011)

Displaying polarization data



Degree of polarization
$p = (Q^2 + U^2)^{1/2}/I$



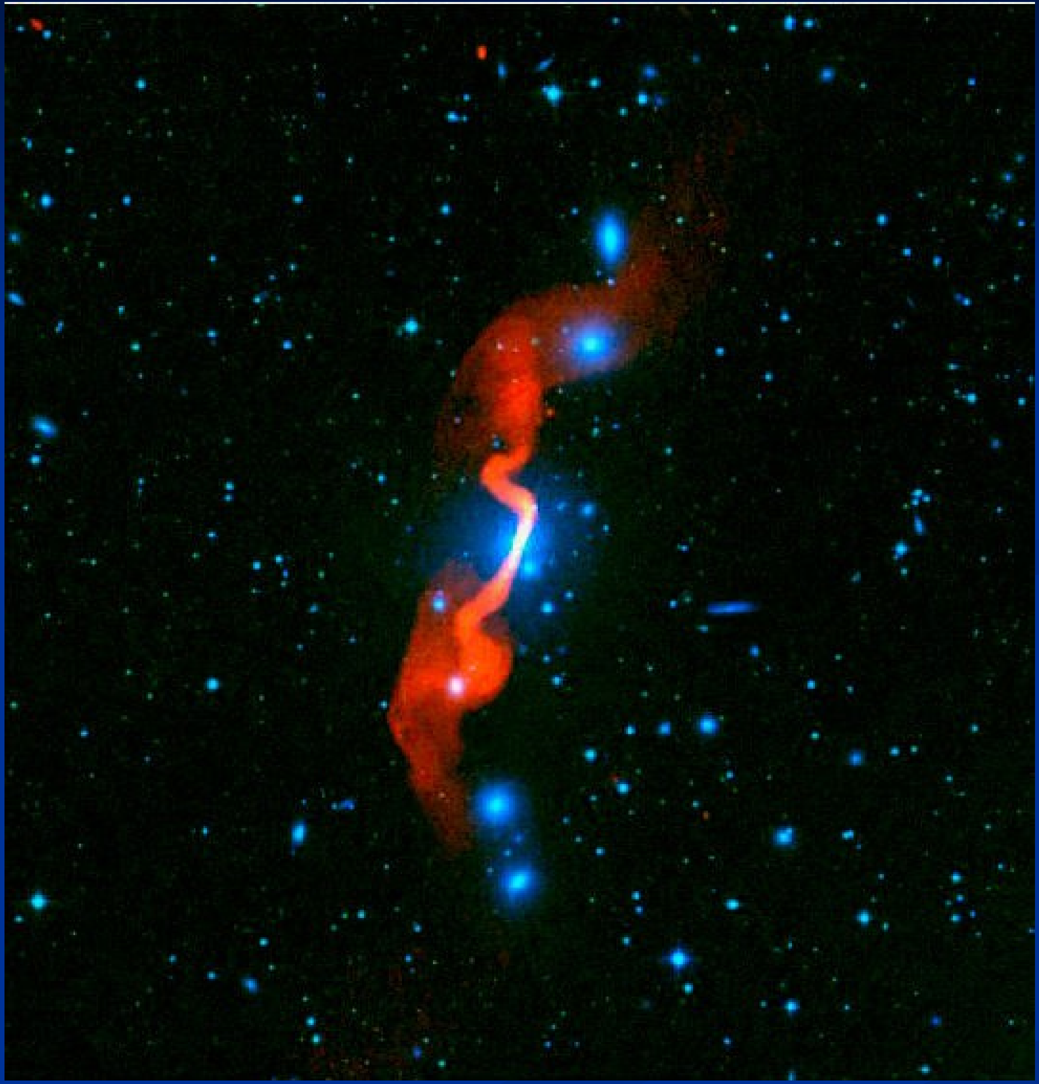
Colour: I

Vectors: PA corrected for Faraday rotation $+90^{\circ}$

Magnitude p

Robert Laing
ESO

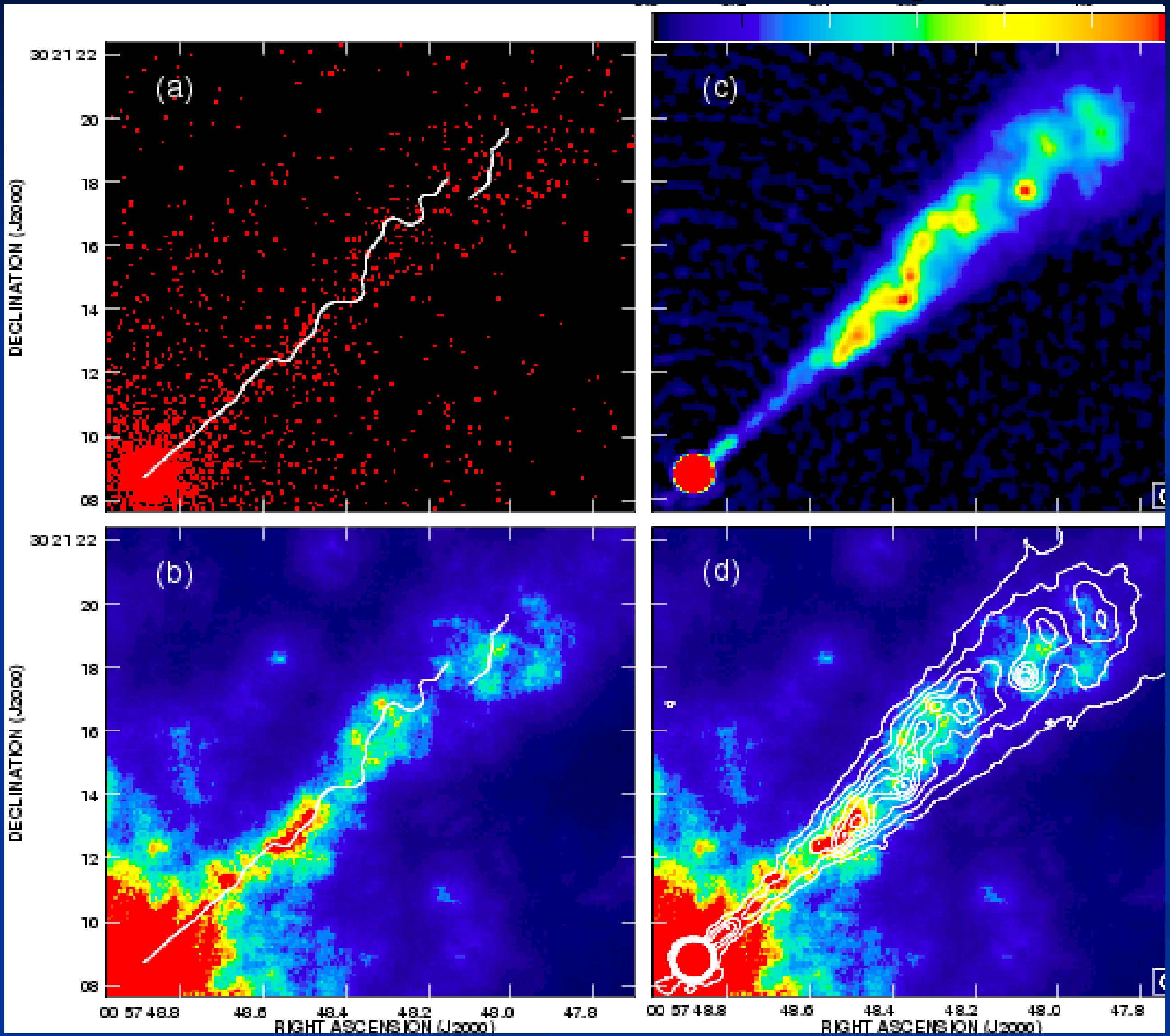
Radio + Optical



Radio Galaxy 3C 31

DSS + VLA

Radio + X-ray



Radio galaxy NGC315

VLA + Chandra

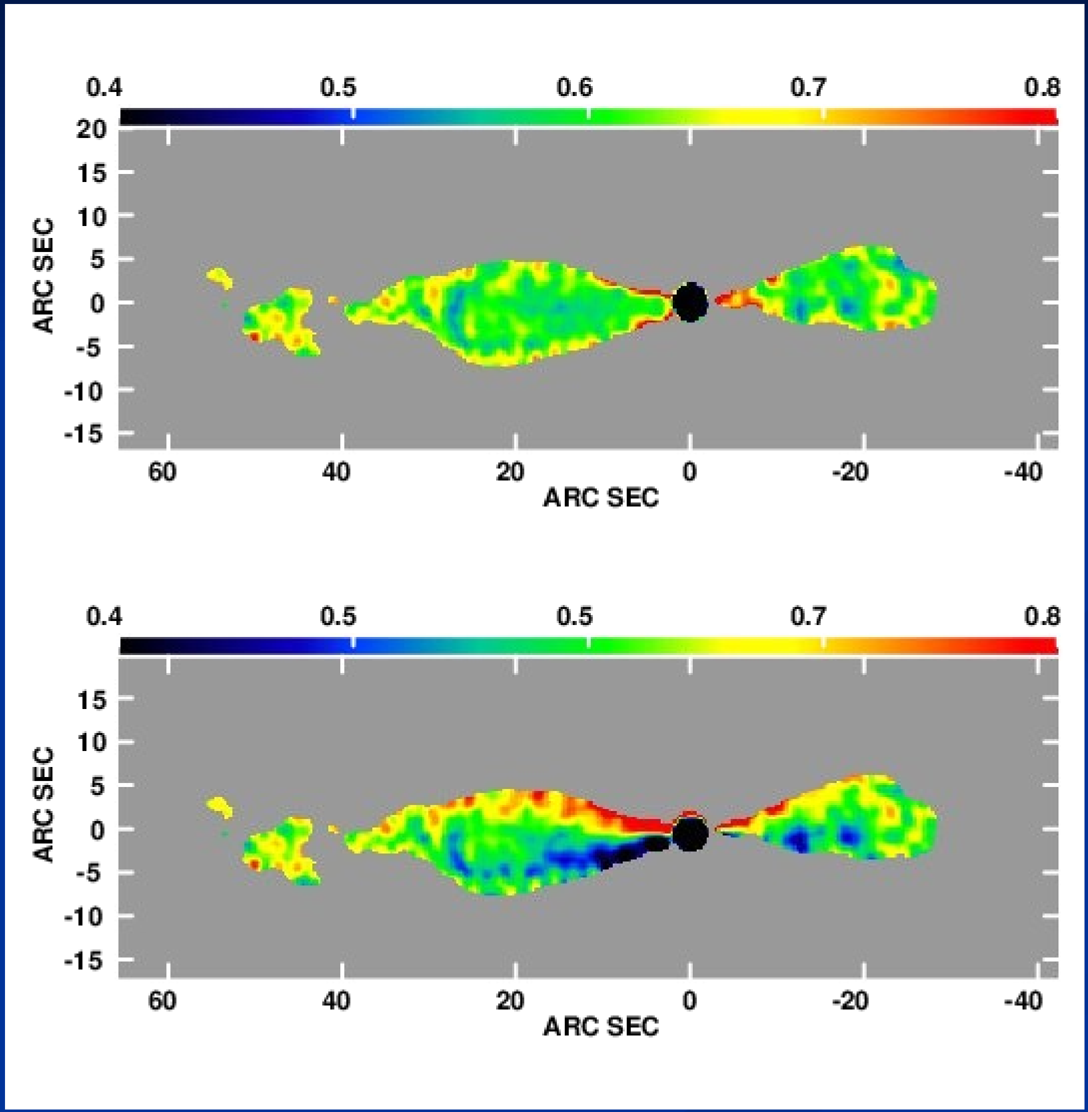
Worrall et al. (2007)

Issues in image registration

- Rationale for image combination
  - Many astrophysical applications require multiwavelength comparison
  - Proper motions may be important
- Regridding
  - Tools available using bilinear or bicubic interpolation (e.g. AIPS OGEOM, OHGEO)
- Accuracy of registration
  - For purely radio data, ideally:
    - calibrator is close to the target
    - use the same phase calibrator for all observations
  - Watch out for errors from ionosphere, troposphere, antenna positions
  - Use internal references if possible (e.g. point sources in the field)
  - Beware changes in structure with frequency
  - N.B.: images at other wavebands may have less accurate absolute astrometry

Good astrometry is vital, but is not the subject of this lecture

Registration Errors

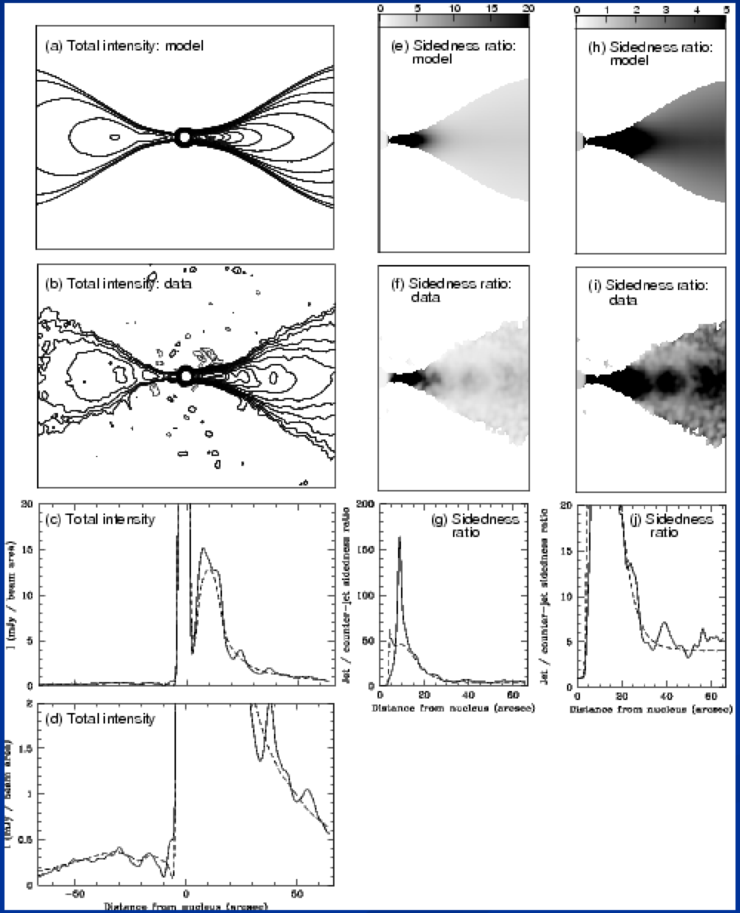


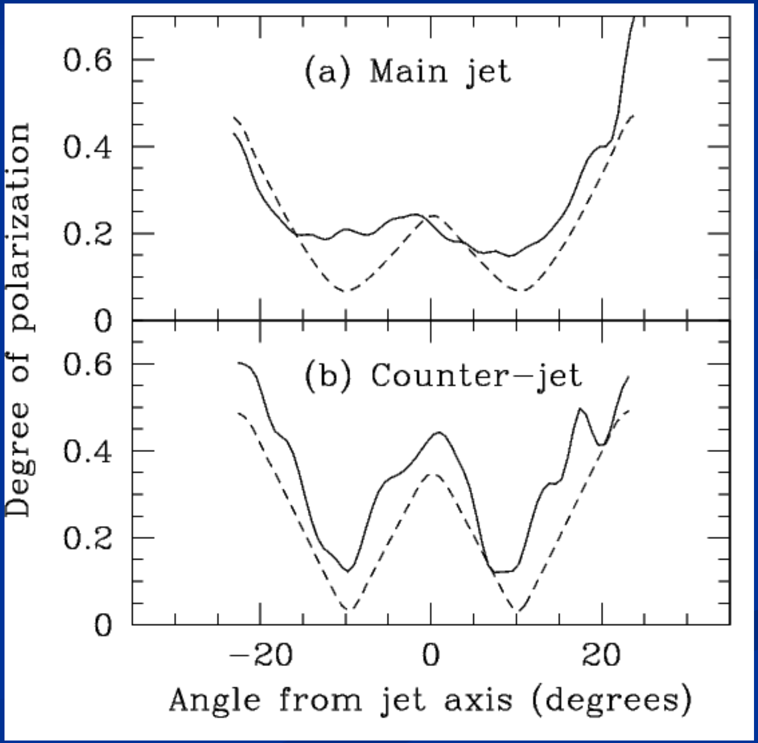
Spectral index image between 1.365 and 4.9 GHZ

Relative shift of 0.2 x FWHM in y direction

Profiles

- 1D profiles across brightness distributions and derived images
- Binning along one direction, or radially





Getting your images where you want them

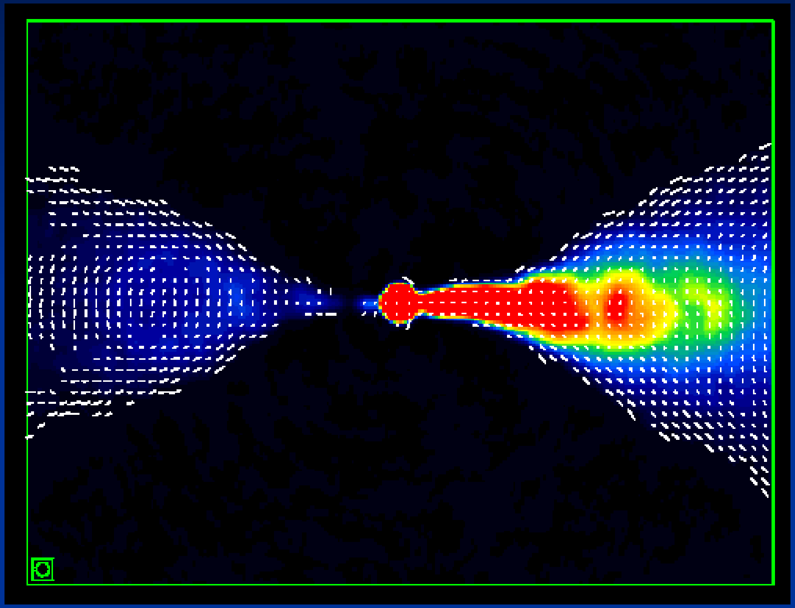
- The FITS standard is a universal interchange format for image data
- It allows you to transfer images between standard packages, but also to read and write images from your own code
- Examples:
  - IDL has a FITS interface
  - cfitsio libraries allow reading and writing of FITS files from C, C++, FORTRAN http://heasarc.gsfc.nasa.gov/fitsio/
  - FORTRAN fragment below reads from a FITS file to an array itemp1

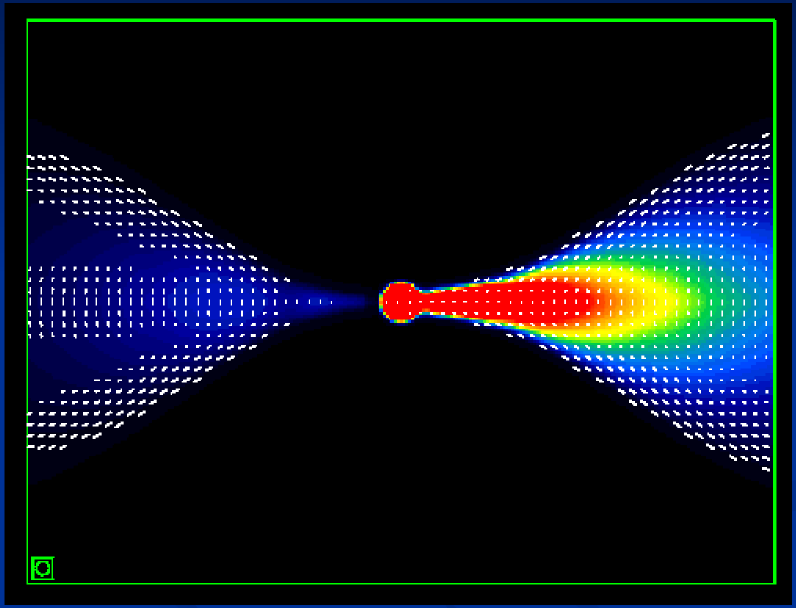
```
*      Get filenames from environment variables

       call getenv ('IMAPFILE1',filename)
       stat = 0
       npix = ((2*mapx1)+1)*((2*mapy1)+1)
       call ftnopn(1,filename,0,stat) ! Open FITS file
       call ftgpve(1,1,1,npix,-99,itemp1,anynull,stat) ! Read FITS image
       call ftclos(1,stat)       ! Close FITS file        to array itemp1
       if (stat .ne. 0) goto 100 ! Error check
```

  - pyFITS/astropy inteface between python and FITS

Modelling: an example





VLA image: I + p vectors;
apparent magnetic field

Model fit; relativistic jet at
38 deg to the line of sight

Canvin et al. (2005)

Summary of image analysis

- Match the resolution to the problem
- For simple images, fit component parameters and derive errors
- Image comparison
  - Simple mathematical operations are easy
  - Regridding and interpolation often required
  - Registration is an issue
  - Noise propagation
- Straightforward to read a FITS image into your own code for more sophisticated modelling

Thanks to Ed Fomalont, from whose NRAO Synthesis Imaging School Lecture I have borrowed extensively and Anita Richards.